

PB86-216793

Applications of Decision Theory  
Techniques in Air Pollution Modeling

(U.S.) Environmental Protection Agency  
Research Triangle Park, NC

Jul 86



U.S. Department of Commerce  
National Technical Information Service

**NTIS**

PB86-216793

EPA/600/3-86/036  
July 1986

APPLICATIONS OF DECISION THEORY TECHNIQUES  
IN AIR POLLUTION MODELING

by

Robert G. Lamb and Saroj K. Hati  
Meteorology and Assessment Division  
Atmospheric Sciences Research Laboratory  
Research Triangle Park, NC 27711

ATMOSPHERIC SCIENCES RESEARCH LABORATORY  
OFFICE OF RESEARCH AND DEVELOPMENT  
U.S. ENVIRONMENTAL PROTECTION AGENCY  
RESEARCH TRIANGLE PARK, NORTH CAROLINA 27711

REPRODUCED BY  
NATIONAL TECHNICAL  
INFORMATION SERVICE  
U.S. DEPARTMENT OF COMMERCE  
SPRINGFIELD, VA. 22161

| TECHNICAL REPORT DATA<br>(Please read Instructions on the reverse before completing)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |                                                   |                                                       |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------|-------------------------------------------------------|
| 1. REPORT NO.<br>EPA/600/3-86/036                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 2.                                                | 3. RECIPIENT'S ACCESSION NO.<br><b>PA86-2167931AS</b> |
| 4. TITLE AND SUBTITLE<br>APPLICATION OF DECISION THEORY TECHNIQUES TO AIR POLLUTION CONTROL                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 5. REPORT DATE<br>July 1986                       |                                                       |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 6. PERFORMING ORGANIZATION CODE                   |                                                       |
| 7. AUTHOR(S)<br>Robert G. Lamb and Saroj K. Hati                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 8. PERFORMING ORGANIZATION REPORT NO.             |                                                       |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br>Same as Block 12.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 10. PROGRAM ELEMENT NO.<br>CDTA1D/03-3049 (FY-86) |                                                       |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 11. CONTRACT/GRANT NO.                            |                                                       |
| 12. SPONSORING AGENCY NAME AND ADDRESS<br>Atmospheric Sciences Research Laboratory -- RTP, NC<br>Office of Research and Development<br>U.S. Environmental Protection Agency<br>Research Triangle Park, North Carolina 27711                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 13. TYPE OF REPORT AND PERIOD COVERED             |                                                       |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 14. SPONSORING AGENCY CODE<br>EPA/600/09          |                                                       |
| 15. SUPPLEMENTARY NOTES                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                   |                                                       |
| 16. ABSTRACT<br><p>This study applies methods of operations research to two basic areas of air pollution modeling: (1) the generation of wind fields for use in models of regional scale transport, diffusion and chemistry; and (2) the application of models in studies of optimal pollution control strategies. The work is illustrated in the context of a hypothetical problem in which optimal sites and emissions control plans are sought for two new power plants.</p> <p>The study addresses all aspects of the problem starting with stochastic specification of the wind fields in the region of interest, proceeding to the development of simple models that relate pollutant emissions to both short and long term averaged concentration, and concluding with the incorporation of game theory concepts into mathematical methods of finding optimum solutions to multi-objective problems. Five objectives are considered in this study: minimization of plant operating cost, and minimization of given short period maximum and long-period averaged concentrations of each of two pollutants. The optimization procedure attempts to fulfill these objectives jointly while maintaining compliance with specified constraints on the overall system.</p> <p>The emphasis of this study is on concepts and technique development rather than the solution of actual problems.</p> |                                                   |                                                       |
| 17. KEY WORDS AND DOCUMENT ANALYSIS                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                                                   |                                                       |
| a. DESCRIPTORS                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | b. IDENTIFIERS/OPEN ENDED TERMS                   | c. COSATI Field/Group                                 |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |                                                   |                                                       |
| 18. DISTRIBUTION STATEMENT<br>RELEASE TO PUBLIC                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 19. SECURITY CLASS (This Report)<br>UNCLASSIFIED  | 21. NO. OF PAGES<br>75                                |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 20. SECURITY CLASS (This page)<br>UNCLASSIFIED    | 22. PRICE<br>\$ 11.95                                 |

# NOTICE

The information in this document has been funded by the United States Environmental Protection Agency. It has been subject to the Agency's peer and administrative review and has been approved for publication as an EPA document. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

## ABSTRACT

This study applies methods of operations research to two basic areas of air pollution modeling: (1) the generation of wind fields for use in models of regional scale transport, diffusion and chemistry; and (2) the application of models in studies of optimal pollution control strategies. The work is illustrated in the context of a hypothetical problem in which optimal sites and emissions control plans are sought for two new power plants.

The study addresses all aspects of the problem starting with stochastic specification of the wind fields in the region of interest, proceeding to the development of simple models that relate pollutant emissions to both short and long term averaged concentration, and concluding with the incorporation of game theory concepts into mathematical methods of finding optimum solutions to multi-objective problems. Five objectives are considered in this study: minimization of plant operating cost, and minimization of given short period maximum and long-period averaged concentrations of each of two pollutants. The optimization procedure attempts to fulfill these objectives jointly while maintaining compliance with specified constraints on the overall system.

The emphasis of this study is on concepts and technique development rather than the solution of actual problems.

## CONTENTS

|                                                                                                                       |      |
|-----------------------------------------------------------------------------------------------------------------------|------|
| Abstract . . . . .                                                                                                    | iii  |
| Figures . . . . .                                                                                                     | vi   |
| Tables . . . . .                                                                                                      | vii  |
| Acknowledgement . . . . .                                                                                             | viii |
| 1. Introduction . . . . .                                                                                             | 1    |
| 2. The Atmospheric transport model . . . . .                                                                          | 8    |
| 3. Derivation of the transport kernels, $P_S$ and $P_L$ . . . . .                                                     | 15   |
| 4. Generating an ensemble of atmospheric flow fields<br>in space-time . . . . .                                       | 29   |
| 5. Optimal solutions to the plant siting problem . . . . .                                                            | 39   |
| 6. Application of the optimization procedure to a<br>hypothetical problem . . . . .                                   | 47   |
| 6.1 Generation of subensembles of flow fields $x_i^n$<br>for each time interval in the simulation period . . . . .    | 48   |
| 6.2 Generation of an ensemble of flows in space-time<br>using the momentum equation and dynamic programming . . . . . | 54   |
| 6.3 Estimation of the dispersion kernels $P_S$ and $P_L$ . . . . .                                                    | 55   |
| 6.4 Determining the optimal plant sites using<br>game theory . . . . .                                                | 56   |
| 7. Conclusion . . . . .                                                                                               | 63   |
| References . . . . .                                                                                                  | 66   |

## FIGURES

| <u>Number</u> |                                                                                                                                                            | <u>Page</u> |
|---------------|------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|
| 1             | Hypothetical example of possible sites (shaded areas) for new power plants and their relationship to existing industry locations and fuel sources. . . . . | 5           |
| 2             | Schematic representation of a sequence of flow fields that constitute a single member of an ensemble of flows in space-time (D,T). . . . .                 | 31          |

# TABLES

| Number |                                                                                                                                                                                                                                                                                                                                                                                                                                                  | Page |
|--------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| 1      | Wind speed and direction observations at hourly intervals at seven surface meteorological stations. A dash (-) in a data column indicates that the value was not reported.                                                                                                                                                                                                                                                                       | 49   |
| 2      | Components of the vector $\underline{w}$ that define the matrix $[Q]$ which in turn defines each member $\underline{x}^n$ of the subensemble of flows at each hour $n$ through Eqs. 3-19 and 3-35. Shown also is the value of the objective function $f(\underline{x})$ defined by (3-25) which is a measure of how well the flow field satisfies the energy spectrum (6-1). The total kinetic energy $E_i$ of the flow is given in column four. | 52   |
| 3      | Optimum design vector $\underline{z}_1^*$ for the extreme problem in which the only objective is the minimization of cost. The optimal vector $\underline{z}_1^*$ and the associated value of the objective function, $F_1(\underline{z}_1^*)$ , are given for two different starting points $\underline{z}_1$ of the search procedure.                                                                                                          | 58   |
| 4      | Optimum design vector $\underline{z}_2^*$ for the extreme problem in which the only objective is minimizing short term $SO_2$ concentration. The optimal vector $\underline{z}_2^*$ and the associated value of the objective function $F_2(\underline{z}_2^*)$ are given for three different starting points $\underline{z}_2$ of the search procedure.                                                                                         | 58   |
| 5      | Optimum design vector $\underline{z}_3^*$ for the extreme problem in which the only objective is minimizing long term $SO_2$ concentration. The optimal vector $\underline{z}_3^*$ and the associated value of the objective function $F_3(\underline{z}_3^*)$ are given for three different initial points $\underline{z}_3$ of the search procedure.                                                                                           | 59   |
| 6      | Optimum design vector $\underline{z}_4^*$ for the extreme problem in which the only objective is minimizing short term particulate concentration. The optimal vector $\underline{z}_4^*$ and the associated value of the objective function $F_4(\underline{z}_4^*)$ are given for three different initial points $\underline{z}_4$ of the search procedure.                                                                                     | 59   |
| 7      | Pareto-optimal design vector $\underline{z}_c^*$ obtained using the game theory concepts presented in Section 5.                                                                                                                                                                                                                                                                                                                                 | 61   |

#### ACKNOWLEDGEMENT

S. K. Hati is greatly indebted to the National Research Council of the National Academy of Sciences and the U.S. Environmental Protection Agency for providing a fellowship that supported this research.

Both authors wish to acknowledge the masterful typing skills of Ms. Barbara Hinton in preparing this document and earlier drafts of it.

## 1. Introduction

Selecting sites for new industries that produce air pollutants is a task that requires crucial trade-offs between capital and operating costs of the industry on the one hand and the requirement to satisfy air quality standards on the other hand. Ideally, one would like to find plant locations at which air quality standards can be met at minimal cost. This study develops a rational procedure for finding such sites. It treats this task as a multi-objective optimization problem subject to given constraints. The objectives it considers are (1) compliance with air quality standards on one or more pollutants, expressed as a limiting long-term average concentration and/or a given short-term average concentration that is not to be exceeded more than once per year; and (2) the total cost of pollution control. In this particular study we focus on the problem of siting power plants, in which case the costs taken into consideration include fuel cost, the cost of transporting fuel from given sources (such as coal mines) to the given plant site and the cost of emissions controls (such as flue gas desulfurization and fly ash precipitation). Optimal tradeoffs among the air quality and cost objectives are sought within the space of feasible solutions using concepts of game theory. Here a "feasible" solution is a plant site, fuel allocation and emissions control design that jointly satisfy all constraints imposed on the system. In the case of power plant siting, constraints include the number and geographical locations of fuel sources; the energy and pollutant contents and the cost of fuel from each source; the required generating capacity of the plant; the geographical areas in which plant construction is admissible; etc.

The basic aspects of the optimal siting problem can be illustrated with the aid of Figure 1. Shown there is a hypothetical situation in which a coal-fired power plant of given generating capacity  $r$  is to be located in a region where polluting industries already exist. The objective is to find a site within the admissible region, indicated by the shaded areas in Figure 1, at which the cost of holding emissions to a level consistent with the allowable air pollution increment is minimum. For a given plant site  $x$ , the total operating costs consists in part of the sum of the costs of the coal purchased from each mine, and the cost of transporting the coal from each mine to the plant site  $x$ . The amount of coal taken from each mine must in turn be compatible with both the supply capacities of each mine and the plant's generating capacity  $r$ , taking into account differences in the energy content of the coal from each mine. The fuel allocation in turn determines the degree of emissions control that the plant must maintain to prevent the effluent concentrations from exceeding the air pollutant increment for that area. We assume that the fly ash and sulfur content of the coal from each mine is known. Within this framework the need for emissions control is handled by defining sources of virtual coal, each of whose ash, sulfur and energy contents is negative and whose cost is given per unit of negative ash or sulfur. The negative energy contents of the virtual coal reflects the fact that sulfur scrubbers and fly ash precipitation consume a portion of the power plants output energy. In the optimization algorithm, the energy requirements of the power plants are met by drawing upon the reserves of real coal and the air quality requirements are satisfied by the addition of virtual coal of either the negative sulfur or negative ash types, or both.

To elaborate briefly, suppose coal is available from two mines with sulfur contents  $s_i$  ( $i = 1, 2$ ), energy contents  $k_i$  (calories/mass) and costs  $c_i$  (dollars/mass). If a given plant consumes coal from each mine at a rate  $x_i$  (mass/time), the uncontrolled emission rate (mass/time) of sulfur would be

$$Q = x_1 s_1 + x_2 s_2$$

and the corresponding energy output and fuel cost would be

$$E = x_1 k_1 + x_2 k_2 \quad ,$$

and

$$C = x_1 c_1 + x_2 c_2 \quad ,$$

respectively. The effect of a sulfur scrubber on the sulfur emission rate  $Q$  can be expressed in the form

$$Q = x_1 s_1 + x_2 s_2 - x_v s_v \tag{1-1}$$

where  $x_v$  is the effective consumption rate (mass/time) of virtual coal whose sulfur content is  $-s_v$  (percent). We can arbitrarily let  $x_v = x_1 + x_2$  in which case  $s_v$  can be expressed in the form

$$s_v = \lambda(x_1 s_1 + x_2 s_2)(x_1 + x_2)^{-1} = \lambda \bar{s} \tag{1-2}$$

where  $\lambda$  is the efficiency of the scrubber ( $0 < \lambda < 1$ ) and  $\bar{s}$  is the percentage sulfur in the uncontrolled plant exhaust. Combining (1-1) and (1-2) we find that the sulfur emission rate of the controlled plant is

$$Q = (1 - \lambda)(x_1 s_1 + x_2 s_2) \tag{1-3}$$

If a scrubber operating at efficiency  $\lambda$  consumes energy at a rate  $k_v(\lambda)$  (calories/mass of coal combustion products processed), the energy output of the controlled plant would be

$$E = x_1 k_1 + x_2 k_2 - (x_1 + x_2) k_v(\lambda) \tag{1-4}$$

and the effective fuel cost would be

$$C = x_1 c_1 + x_2 c_2 + (x_1 + x_2) c_v(\lambda) \quad , \tag{1-5}$$

where  $c_v(\lambda)$  is the operating cost (dollars) of a scrubber of efficiency  $\lambda$  per mass of coal combustion products processed. Since  $k_v(\lambda)$  and  $c_v(\lambda)$  are in principle known functions, both approaching zero as  $\lambda \rightarrow 0$  and increasing, perhaps exponentially as  $\lambda \rightarrow 1$ , the plant's output energy and effective fuel cost is set once we have selected a value for  $\lambda$  that allows us to achieve a given sulfur emissions reduction.

For a given allocation of fuel from the mines to the plant the uncontrolled emissions of the plant can be estimated, and with this knowledge and suitable atmospheric transport models, the impact of the plant on air quality can be determined. Comparing the predicted pollutant concentrations with the allowed pollutant increments determines the degree of emissions control that must be imposed, and this in turn establishes the total emissions control necessary and hence the operating cost of the plant at the given site. Thus, the problem of optimal plant siting requires an iterative manipulation of the design variables, namely fuel allocation, emissions control and plant location, until the configuration of least cost is found. Two essential elements are required to implement this process: (1) an atmospheric transport model that can predict the air quality impact of given plant emissions, and (2) a rational procedure for manipulating the design variables. The latter is a critical component of any optimization scheme because without an efficient means of scanning the phase space of the design variables, the solutions produced may be artifacts of the scheme, rather than optimal points; and/or the computer time required to generate them may be prohibitive. The method we use in the present study is a mathematically sophisticated procedure that is

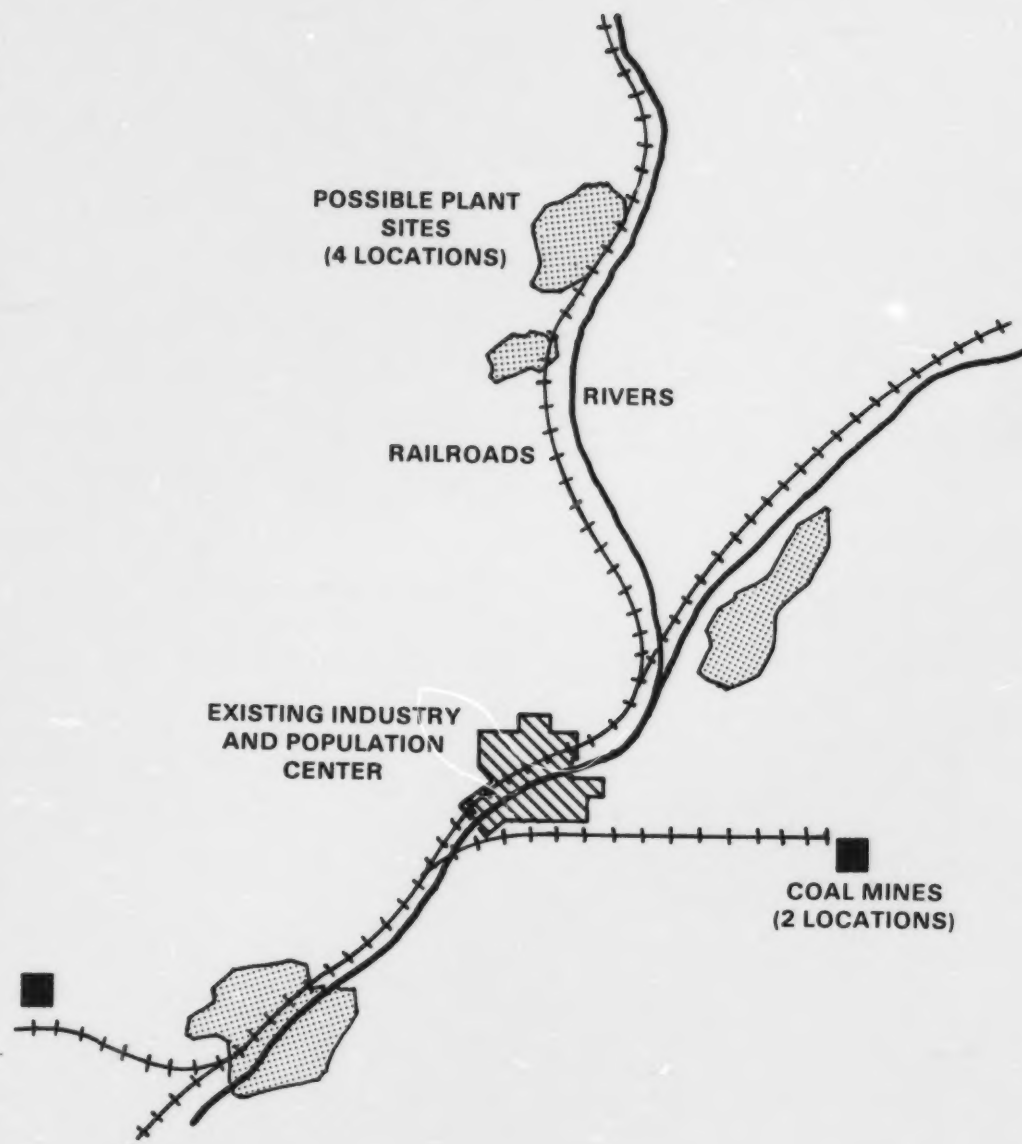


Figure 1. Hypothetical example of possible sites (shaded areas) for new power plants and their relationship to existing industry locations and fuel sources.

based on the concepts of game theory. It yields solutions in which the design objectives are optimized jointly. That is, each objective is satisfied to the fullest extent possible without unduly restricting the degree to which each of the other objectives can be met.

In this regard it is important to realize that under current U.S. air pollution regulations, a new industry that is to be located in a region where air quality standards are in attainment is forbidden to emit pollutants at a rate that would cause concentrations to exceed the standard. In this case the pollutant concentrations that the industry would produce become a part of the constraints on the system design rather than a variable that is to be minimized. As far as our optimization procedure is concerned, this is a trivial situation to treat and consequently it does not make a particularly good case to consider for demonstration purposes. Therefore, in this report we will focus on problems in which the air quality impact is a variable to be minimized jointly with cost, rather than a variable that is limited by constraints. Concerning the atmospheric transport model component of our optimization procedure, we develop a new model in Section 3 and 4 based on ideas presented recently in Lamb (1983, 84) in which the long-range transport of pollutants is treated mathematically as a subset of the solution space of an underdetermined system of equations.

This study is not the first attempt to apply optimization techniques to problems of air pollution control. Seinfeld and his co-workers have reported a number of efforts along this line. Seinfeld and Kyan (1971) considered the problem of finding the emissions control strategy for sources in a given airshed that would yield a given level of air quality

at least cost. This work was extended by Kyan and Seinfeld (1972) to treat the case of control plans that are implemented over a period of years. Again, the objective was to find the least cost (multiyear) emissions control strategy that would maintain a given level of air quality, taking into account that the cost of controlling a given source in a given year is a function of the type of control equipment that was installed in previous years. The authors utilize dynamic programming in their procedure, which is one of the techniques we shall employ here.

The major difference between the present study and the earlier studies with which we are familiar is our consideration of multiple objectives in the optimization process. That is, instead of considering the cost of emissions controls as the only variable to be minimized, subject to constraints imposed by air quality standards, we consider the control costs and pollutant concentrations as a set of variables that are to be minimized jointly. In this task we make use of game theory concepts. One of the advantages of our approach is that it reveals the bounds on each of the objective variables that are achievable when any one of the variables is minimized alone. The information provided by these bounding values is potentially very valuable in determining where the limits of feasible air quality standards lie.

Following a brief discussion in the next section of the basic equations for short and long-term averaged concentration, we develop the transport model component of our optimization procedure in Section 3 and 4, and the optimization technique itself in Section 5. An elementary application of the overall method is illustrated in Section 6.

## 2. The Atmospheric Transport Model

The transport model provides the mathematical link between the given emissions of a source and its resulting impact on air quality. The National Air Quality Standards which have been established in the United States for pollution control are usually expressed either as an annual arithmetic mean value of the pollutant or as maximum 24 hour average concentration that is not to be exceeded more than once a year. These are the two measures of environmental impact that we would like to relate to source emissions.

Lamb (1976), referred to hereafter as L76, developed mathematical expressions for both the long-term mean concentration and its higher order moments, from which estimates of the peak short-term averaged concentration can be determined. His expression for the annual averaged concentration at a receptor located at point  $\underline{r}$  at time-of-day  $\xi$ , where  $0 \leq \xi < 24$  hour is

$$\langle c(\underline{r}, \xi) \rangle = \iint_{-\infty}^{\xi} p(\underline{r}, \xi | \underline{r}', t') S(\underline{r}', t') dt' dr' \quad (2-1)$$

where

$$S(\underline{r}', t') = \text{source strength at point } \underline{r}' \text{ at time-of-day } t', \quad (2-2)$$

and it is assumed that  $S$  has diurnal periodicity, i.e.,

$$S(\underline{r}', t') = S(\underline{r}', t' + 24 \text{ hr})$$

and  $0 \leq t' < 24$  hour. Also, in (2-1) the kernel  $p$  is the probability density that a particle located at  $\underline{r}$  at time of day  $\xi$  was located at the point  $\underline{r}'$  at  $(\xi - t')$  hours earlier. This kernel is a Lagrangian, climatological property of the atmosphere in the region of the source and receptor points  $\underline{r}'$  and  $\underline{r}$ .

The annual averaged concentration at  $\underline{r}$ , that is the average taken over the entire year, is found from (2-1) to be

$$\langle c(\underline{r}) \rangle = \frac{1}{24} \int_0^{24} \int_{-\infty}^{\xi} p(\underline{r}, \xi | \underline{r}', t') S(\underline{r}', t') dt' d\underline{r}' d\xi . \quad (2-3)$$

Keeping in mind that  $p$  is a climatological property of the given region, which we shall assume is known, we see that (2-3) provides a means of determining the impact that a given distribution of sources  $S$ , which is one of the design variables in our optimization problem, will have on annual averaged concentrations of pollutants anywhere  $\underline{r}$  in the region.

It is also shown in L76 that the probability that the concentration at a given receptor  $\underline{r}$  will exceed a given value  $c_m$ , say, for a given length of time, say  $\tau$ , beginning at time-of-day  $\xi$  is given by

$$\text{prob} [c(\underline{r}, t) > c_m \text{ for } \xi \leq t \leq \xi + \tau] \leq \frac{R(\tau; \underline{r}, \xi)}{c_m^2} \quad (2-4)$$

where

$$R(\tau; \underline{r}, \xi) = \iiint_{-\infty}^{\xi} \int_{-\infty}^{\xi + \tau} p_2(\underline{r}, \xi; \underline{r}, \xi + \tau | \underline{r}', t', \underline{r}'', t'') S(\underline{r}', t') S(\underline{r}'', t'') dt' dt'' d\underline{r}' d\underline{r}'' \quad (2-5)$$

where the kernel  $p_2$  is another (known) climatological property of the given region. If the air quality standard stipulated that the hourly averaged concentration could not exceed a value  $\Gamma$  more than once a year, then by casting the "once a year" condition into a probability equivalent and setting  $\tau = 1$  hour and  $c_m = \Gamma$ , we would have in (2-4) and (2-5) a quantitative condition that our source distribution  $S$  would have to meet to comply with the standards. Thus, equations (2-3)-(2-5) provide a

link between the pollutant emissions and the resulting air quality, given the climatological properties of the region in the Lagrangian forms  $p$  and  $p_2$ .

In this report our principal concern is with the development of mathematical techniques for solving the optimization problem that we posed in the Introduction. In order to expedite that development and to facilitate demonstrations of the resulting procedure, we will express the relationships embodied in (2-3) and (2-4) in a primitive fashion that will form a simple operational basis for deriving the short and long-term averaged concentration estimates needed in the optimization scheme. Evaluating (2-3) and (2-4) rigorously would be necessary for any formal application of this work but it is far beyond the needs or scope of the present technique development effort.

Confining ourselves to two dimensions, namely the horizontal plane, and letting

$$\Delta R = (\Delta x, \Delta y) \quad (2-6)$$

denote the dimensions of the sampling volume centered at a receptor point  $\underline{r}$ , the material concentration  $c$  at  $\underline{r}$  at time  $t$  can be expressed by (following the procedure used in L76 to derive eqn. 2-3 above)

$$c(\underline{r}, t) = \frac{1}{\Delta x \Delta y} \sum_{n=1}^N p_n(\underline{r}, t'; \Delta R) \quad (2-7)$$

where concentration  $c$  is defined as the number of marked (pollutant) particles per volume of fluid, and  $N$  is the total number of such particles released by all sources up to the time  $t$ . The kernel  $p_n$  in (2-7) is defined as

$$p_n(\underline{r}, t'; \Delta R) = \begin{cases} 1, & \text{if particle } n \text{ is located inside } \Delta R \text{ centered} \\ & \text{at } \underline{r} \text{ at time } t; \\ 0, & \text{otherwise} \end{cases} \quad (2-8)$$

Clearly  $p_n$  is a function of the fluid velocity field. We will consider it in detail in the next Section.

Defining a moving time average of concentration over an arbitrary period  $T$  to be

$$\bar{c}^T(\underline{r}, t) \equiv \frac{1}{T} \int_{t - \frac{T}{2}}^{t + \frac{T}{2}} c(\underline{r}, t') dt' \quad (2-9)$$

we obtain upon substituting (2-7) into (2-9) and using (2-8)

$$\bar{c}^T(\underline{r}, t) = \frac{1}{T \Delta x \Delta y} \sum_{n=1}^N \Delta t_n(\underline{r}, t; \Delta R) \quad (2-10)$$

where

$$\Delta t_n(\underline{r}, t; \Delta R) = \begin{aligned} &\text{total time spent by particle } n \\ &\text{inside } \Delta R \text{ centered at } \underline{r} \text{ during} \\ &\text{the period } (t - \frac{T}{2}) < t' < (t + \frac{T}{2}). \end{aligned} \quad (2-11)$$

Eq (2-10) provides a convenient way of computing both short and long-term average concentration. However, it is important to note that  $\Delta t_n$ , and hence the time averaged concentration  $\bar{c}^T$ , is an implicit function of the points of origin of the  $N$  particles. For example, for a given receptor site  $\underline{r}$  and measurement time  $t$  it is obvious that the value of  $\Delta t_n$  will depend on the release point of particle  $n$  in space-time relative to  $(\underline{r}, t)$ . We will make the dependence of  $\bar{c}^T$  on the particle release locations  $\underline{x}_N = (\underline{x}_1, \underline{x}_2, \dots, \underline{x}_N)$

explicit by writing  $\bar{c}^T(\underline{r}, t; \underline{x}_N)$ .

We are interested in two measures of air quality impact, namely concentration averaged over a period  $T = T_S = 1$  day and concentration averaged over a year, i.e.,  $T = T_L = 1$  year. In particular, we will use as a measure of the emissions impact the maximum values of these averages over the spatial domain containing the sources and over a one year period, specifically

$$C_S(\underline{x}_N) = \max_{\substack{\underline{r} \in D \\ t \in T_L}} \{ \bar{c}^{T_S}(\underline{r}, t; \underline{x}_N) \} \quad (2-12)$$

$$C_L(\underline{x}_N) = \max_{\underline{r} \in D} \{ \bar{c}^{T_L}(\underline{r}; \underline{x}_N) \} \quad (2-13)$$

where  $T_S$  and  $T_L$  are 1 day and 1 year, respectively, and  $D$  is the spatial domain of interest. In the optimization procedure that we develop later, locations  $\underline{x}_N$  for the sources are sought that minimize  $C_S$  and  $C_L$  jointly with cost.

In the example problem that we solve later in Section 6, we consider the case where power plants of equal emission rates are to be sited. In this case each of the sources is treated as a point, one at  $\underline{x}_{S1}$ , another at  $\underline{x}_{S2}$ , and so on where the  $\underline{x}_{Sn}$ ,  $n=1, \dots, N$ , are variables to be determined by the optimization procedure. Furthermore, the strengths of each power plant, which are equal, become multiplicative factors in (2-10) with the result that we can express (2-12) and (2-13) in the simple forms

$$C_S(\underline{x}_N) = Q \cdot \max_{t \in T_L} \{ P_S(t; \underline{x}_{S1}, \dots, \underline{x}_{SN}) \} \quad (2-14)$$

$$C_L(\underline{x}_N) = Q \cdot P_L(\underline{x}_{S1}, \dots, \underline{x}_{SN}) \quad (2-15)$$

where  $Q$  is the emission rate of each source;  $P_S$  is the maximum short-term ( $T_S = 1$  day) concentration at time  $t$  in the domain  $D$  due to  $N$  point sources of unit strengths located at  $\underline{x}_{Sn}$ ,  $n=1,2,\dots,N$ ; and  $P_L$  is the maximum long-term ( $T_L = 1$  year) concentration in  $D$  due to the same unit sources.

In order to use (2-14) and (2-15) in the optimization procedure that we develop later, it is convenient to express  $P_S$  and  $P_L$  as polynomials in the  $2N$  variables  $\underline{x}_{Sn}$ . For example, in the case where  $N=2$  and  $(\underline{x}_{S1}, \underline{x}_{S2}) = (x_1, y_1, x_2, y_2)$  we can write

$$P_S(\underline{x}_{S1}, \underline{x}_{S2}) = a_1 + a_2x_1 + a_3y_1 + a_4x_2 + a_5y_2 + a_6x_1^2 + \dots a_{18}x_2y_2^2 + a_{19}y_2^3 \quad (2-16)$$

where the coefficients  $a_i$  are functions of  $t$ . By evaluating  $P_S$  for a number of combinations of source sites  $\underline{x}_{S1}, \underline{x}_{S2}$ , in particular a number larger than the number of coefficients in (2-16), we can evaluate the  $a_i$  by least squares approximation methods. We will use this technique later when we demonstrate an application of the optimization procedure.

Going back to Eqs. (2-14 and 2-15) for the short and long-term mean concentrations, and comparing these expressions with (2-12) and (2-13), we see that the kernels  $P_S$  and  $P_L$  are directly derivable from Eq. (2-10). That equation is in turn a function of the time  $\Delta t_n$  that the  $n$ -th pollutant particle spends in the receptor volume  $\Delta R$  centered at the receptor site  $(\underline{r}, t)$ ; and  $\Delta t_n$  is determined in turn solely by the wind

fields in the domain of interest. Specifically, the position coordinates of the n-th particle at time t are

$$x_n(t) = x_{n0} + \int_{t_0}^t u[x_n(t'), y_n(t'), t'] dt' \quad (2-17a)$$

$$y_n(t) = y_{n0} + \int_{t_0}^t v[x_n(t'), y_n(t'), t'] dt' \quad (2-17b)$$

where  $(u,v)$  denotes the horizontal wind components. The value of  $\Delta t_n$  is simply the total time that  $(x_n, y_n) \in \Delta R$ , where  $\Delta R$  is the sample volume centered at  $(\underline{r}, t)$ . Thus, Eqs. (2-17) and (2-10) form the basis for computing the dispersion kernels  $P_S$  and  $P_L$ . In the next two sections we will derive functional descriptions of the wind that will permit evaluation of the trajectory equations (2-17). In so doing we will show that the atmospheric flow field can be specified only to within a set or ensemble of functions  $(u_i, v_i)$ ,  $1 < i < \infty$ . We can give a likelihood or probability that any particular member  $(u_i, v_i)$  of the flow ensemble is the correct description in a given instance but we cannot provide a full description of the actual flow with complete certainty. With the winds specified only to within an ensemble of functions, we cannot specify the pollutant concentrations that will arise from a given set of sources with certainty. In under-determined problems of this type we often work with "expected" values, which are defined as probability weighted averages over the ensemble. Consequently, when we refer from now on to the kernel functions  $P_S$  and  $P_L$  or to the associated short and long-term concentration values  $C_S$  and  $C_L$ , we refer to the "expectation" of these quantities. More on this in the next Section.

### 3. Derivation of the transport kernels, $P_S$ and $P_L$

Earlier in Eq. (2-7) we expressed material concentration  $c$  in terms of a particle marker function  $p_n$ , which we defined in (2-8) to have a value of either 1 or 0 depending on whether or not the  $n$ -th material particle happens to fall within the sampling volume of the concentration measurement. (Here we use the term particle to refer to either gas molecules or aerosols.) Since the motion of particles in the atmosphere is determined almost exclusively by the atmospheric velocity field, deriving quantitative estimates of the marker function  $p_n$ , and hence the subsequent kernel functions  $P_S$  and  $P_L$ , requires detail knowledge of the winds.

The conventional approach to describing atmospheric flow in diffusion studies has been to represent the wind velocity at any point as the sum of two parts, a so-called mean component which is associated with the phenomenon of transport and a random part which represents the turbulent velocity fluctuations that give rise to the phenomenon of diffusion. The mean component is treated as a deterministic function that can be derived from meteorological wind data. By contrast the turbulent wind component is considered to be a stochastic variable whose effects on material spread must be parameterized in terms of known or measurable statistical properties of the turbulent velocity. Representing the winds in terms of a mean and random part in air pollution is a technique that was borrowed from the early engineering studies of fluid motion in pipes and over flat surfaces. In small scale flows of these types the concept of a mean motion is well defined because if one performs moving time averages of

the instantaneous velocity at a fixed point inside a pipe or in a boundary layer over a flat surface, the average velocity tends to a constant value as the averaging time is increased. (This is true, of course, only if the forces driving the flow are steady in time.) In this case the turbulent component of the velocity is defined as the difference between the mean part and the instantaneous velocity. Averaging moments of the turbulent velocity over time yields so-called turbulence statistics which may be used, among other things, to estimate the rate of spread of material released into such flows. Early investigators of atmospheric diffusion found that these concepts provided a viable basis for estimating material dispersion over limited distances, and thus they were incorporated into early models such as the so-called Gaussian plume formula. This expression provides estimates of time averaged concentration of material from a steady point source in terms of three parameters: a mean wind speed  $\bar{u}$ , which one can estimate using wind sensors located near the source, and two parameters  $\sigma_x$  and  $\sigma_y$  which represent the effects of the turbulent wind fluctuations on material spread. The last two parameters are treated as universal functions of  $\bar{u}$ , the surface heat flux, and the surface roughness, all of which are measurable parameters.

Microscale models such as the Gaussian formula are not adequate for power plant siting studies because those models are valid only within distances of a few kilometers from a source. A model capable of simulating transport and diffusion over spatial scales of 100 to 1000 km and periods of several days is needed. A fundamental question concerning the modeling of large scale domains is how meteorological data should be used to describe the flow field. The assumptions of stationarity and homogeneity

that are implicit in definitions of the mean and turbulent wind components used in microscale models are clearly not valid over large time and space domains. Nevertheless, large scale transport models have been developed in recent years in which the transport flow is obtained simply by applying ad hoc interpolation schemes to the set of microscale mean wind values measured at each wind station.

Lamb [1983, '84] studied the problem of representing the flow field in regional scale (1000 km) models and concluded that on that scale one cannot define a deterministic component of the flow that is analogous to the so-called mean wind used in microscale models. He argued that on the regional scale meteorological data and physical principles together are capable of delineating only a set of flow fields any one of which is a possible description of the flow that actually existed in the region during the period the wind data were collected. One can assign weights, or likelihood estimates to each flow field in the set of possible flows and by this process arrive at a finite set or ensemble of flow fields each of which has a nonzero probability of occurrence. The weights one assigns to each member of the set of possible flows is based on additional information that one might have available, such as climatological wind statistics for the region of interest. In any event one ends up with an ensemble of flow fields, rather than a single field as the conventional approach would yield.

In the remainder of this Section and in the Section that follows, we describe the flow field ensemble approach in some detail and we develop an implementation of it that is suitable for generating the particle marker function  $p_n$  and the transport kernels  $P_S$  and  $P_L$  introduced in the previous Section.

Let us make clear at the outset that by a "flow field" we mean a function pair  $[u(x,y,t), v(x,y,t)]$  that describes the x-y components of the atmospheric motion at each point  $(x,y)$  in a domain  $D$ , which is generally an area of order 1000 km on a side, and at each instant  $t$  in the interval  $0 < t < T$ , where  $T$  may be several days or longer. By an "ensemble" of flows, we mean a set of  $I$  functions

$$[u(x,y,t), v(x,y,t)]_i, \quad i = 1 \dots I \quad (3-1)$$

in which the  $i$ -th member is different than all the other  $(I-1)$  members.

Each member of the flow ensemble (3-1) uniquely determines a marker function  $p_n$ , defined earlier in (2-8), and hence also kernel functions  $P_S$  and  $P_L$  as well. Thus, there exists an ensemble of  $P_S$  functions and one of  $P_L$  kernels in which each member has a one-to-one correspondence with a member of the flow ensemble. And it follows from this and (2-14) and (2-15) that there exist also ensembles of short and long-term peak concentration  $C_S$  and  $C_L$  each member of which is linked to a member of the flow ensemble (3-1). If we assign a probability  $\rho_i$  to each member of (3-1), such that

$$\sum_{i=1}^I \rho_i = 1, \quad (3-2)$$

the value  $\rho_i$  applies also to the members of the  $p_n$ ,  $P_S$  and  $P_L$  ensembles that are associated with  $[u,v]_i$ . With a number  $\rho_i$  that we are calling a probability or likelihood associated with each member of each ensemble, we extend our statistical interpretation of our as yet undefined ensemble further by defining the "expectation" of any property  $\zeta$  of an ensemble to be

$$\langle \zeta \rangle \equiv \frac{1}{I} \sum_{i=1}^I \zeta_i \rho_i \quad (3-3)$$

where the angle brackets  $\langle \rangle$  denote an ensemble average. Thus, using (3-3) and (2-7), we define the "expected" concentration at any point  $(\underline{r}, t)$  in  $(D, T)$  to be

$$\langle c(\underline{r}, t) \rangle = \frac{1}{\Delta x \Delta y} \sum_{n=1}^N \langle p_n(\underline{r}, t; \Delta R) \rangle \quad (3-4)$$

where the marker function expectation is given by

$$\langle p_n \rangle = \frac{1}{I} \sum_{i=1}^I p_i p_n(i) \quad (3-5)$$

In a similar fashion we can define the expected peak short and long-term concentrations  $\langle C_S \rangle$  and  $\langle C_L \rangle$ , respectively:

$$\langle C_S(\underline{x}_N) \rangle = Q \cdot \max_{t \in T_L} \{ \langle P_S(t; \underline{x}_{S1}, \underline{x}_{S2}) \rangle \} \quad (3-6)$$

$$\langle C_L(\underline{x}_N) \rangle = Q \langle P_L(\underline{x}_{S1}, \underline{x}_{S2}) \rangle \quad (3-7)$$

where  $\langle P_S \rangle$  and  $\langle P_L \rangle$  have definitions similar to (3-5). It is the expectation of concentration variables such as  $\langle C_S \rangle$  rather than any single member  $C_{Sj}$  of the associated ensemble that one should use in any air quality planning studies, such as the power plant siting problem considered here. In order to implement this approach we must first define the flow ensemble (3-1) and the associated set of probabilities  $p_i$ . This is the task we undertake in the remainder of this Section and in Section 4. Once these steps are complete, we will have arrived at formulations of  $\langle C_S \rangle$  and  $\langle C_L \rangle$  that we can use in the optimal siting procedure.

The flow ensemble is defined simplest in terms of a Fourier expansion of the wind components  $u$  and  $v$ , namely

$$u(x,y,t) = \sum_{k=-M}^M \sum_{l=-M}^M U(k,l;t) e^{i(k\mu x + l\nu y)} \quad (3-8)$$

$$v(x,y,t) = \sum_{k=-M}^M \sum_{l=-M}^M V(k,l;t) e^{i(k\mu x + l\nu y)} \quad (3-9)$$

where  $M$  is a given integer and

$$\mu = 2\pi/L_x \quad (3-10a)$$

$$\nu = 2\pi/L_y \quad (3-10b)$$

and  $(L_x, L_y)$  are the dimensions of the simulation domain  $D$ , or larger. In (3-8) and (3-9)  $U$  and  $V$  are complex Fourier amplitudes. For convenience we will drop the reference to the time dependence of these variables wherever it is unimportant in the analyses. Since  $u$  and  $v$  are real variables, we find from (3-8, 9) that  $U$  and  $V$  must satisfy the conditions

$$U^*(k,l) = U(-k,-l) \quad (3-11a)$$

$$V^*(k,l) = V(-k,-l) \quad (3-11b)$$

where  $*$  denotes the complex conjugate. In terms of their real and complex parts, relations (3-11) can be written

$$U_r(k,l) = U_r(-k,-l) \quad (3-12a)$$

$$U_i(k,l) = -U_i(-k,-l) \quad (3-12b)$$

and similarly for  $V$ . Here

$$U(k,l) = U_r(k,l) + iU_i(k,l) \quad (3-13)$$

A similar expression relates  $V$ ,  $V_r$  and  $V_i$ .

In order for a given mathematical function to describe atmospheric motion, it must be consistent with the equations that describe the so-called physical laws that govern fluid motion. One of these laws is the

conservation of mass which, in the case of a shallow, 2-D fluid is expressed by

$$\partial u / \partial x + \partial v / \partial y = \delta(x,y) \quad (3-14)$$

where  $\delta$  is the given horizontal divergence of the flow. Taking the Fourier transform of (3-14) we get

$$ik_{\mu}U + il_{\nu}V = D \quad (3-15)$$

where  $D(k,l)$  is the transform of the given function  $\delta(x,y)$ . Eq. (3-15) imposes the following relationship on the transform amplitudes:

$$k_{\mu}U_j + l_{\nu}V_j = -D_r \quad (3-16a)$$

$$k_{\mu}U_r + l_{\nu}V_r = D_j \quad (3-16b)$$

Thus in searching for mathematical descriptions of the horizontal winds, we restrict ourselves to transforms  $[U(k,l), V(k,l)]$  that satisfy (3-16).

For the purposes of the present study we will consider only nondivergent flows, i.e.,  $\delta = 0$ , in which case (3-16) gives

$$V_r(k,l) = -\left(\frac{k_{\mu}}{l_{\nu}}\right) U_r(k,l) \quad , \quad l \neq 0; \quad (3-17a)$$

$$V_j(k,l) = -\left(\frac{k_{\mu}}{l_{\nu}}\right) U_j(k,l) \quad , \quad l \neq 0. \quad (3-17b)$$

and

$$U_r(k,0) = U_j(k,0) = 0 \quad , \quad l=0, \quad k>0. \quad (3-17c)$$

These conditions significantly reduce the number of independent Fourier variables that describe nondivergent, horizontal atmospheric motion.

The law of mass conservation (3-14) is but one constraint on the set of functions that describe atmospheric flow. Another constraint is compatibility with the observed wind velocities in the domain of interest.

In particular, we must have

$$\sum_{k=-M}^M \sum_{l=-M}^M U(k,l) e^{i(k\mu x_j + l\nu y_j)} = \hat{u}_j \quad (3-18)$$

$$\sum_{k=-M}^M \sum_{l=-M}^M V(k,l) e^{i(k\mu x_j + l\nu y_j)} = \hat{v}_j \quad (3-19)$$

$$j = 1, \dots, J.$$

where  $(\hat{u}_j, \hat{v}_j)$  is the wind velocity measured at point  $(x,y) = (x_j, y_j)$ . The value  $(\hat{u}_j, \hat{v}_j)$  applies to a given time  $t_1$  and we assume that a total of  $J$  measurements were made at this time at given sites  $(x_j, y_j)$ ,  $j = 1, \dots, J$ . The number  $J$  and locations of the measurements may change as  $t$  changes. On expanding (3-18) and (3-19) and making use of conditions (3-12) and its  $V$  counterpart, and constraint (3-17), we find that the flow field descriptions we seek are contained in the set of transforms  $(U,V)$  that satisfy the following system of equations:

$$\sum_{k=-M}^M \sum_{l=-M}^M [a_j(k,l)U_r(k,l) + b_j(k,l)U_i(k,l)] + U_r(0,0) = \hat{u}_j \quad j=1, \dots, J$$

$$\sum_{k=-M}^M \sum_{l=1}^M [c_j(k,l)U_r(k,l) + d_j(k,l)U_i(k,l)] \quad (3-20)$$

$$+ \sum_{k=1}^M [e_j(k)V_r(k,0) + f_j(k)V_i(k,0)] + V_r(0,0) = \hat{v}_j$$

where  $a_j, b_j, \dots, f_j$  are functions of  $(k, l)$  and the locations  $(x_j, y_j)$  of the observation stations. Note, too, that due to condition (3-12b) the imaginary parts of  $U(0,0)$  and  $V(0,0)$  are identically zero. The constant terms  $U_r(0,0)$  and  $V_r(0,0)$  are given.

The equation set (3-20) contains  $2J$  equations in  $4M(M+1)$  unknowns. Recall that  $J$  is the number of meteorological stations reporting at a given hour and is therefore a fixed number, while  $M$  is arbitrary inasmuch as it is the total number of wave number amplitudes in the Fourier half-space (see 3-8,9) that we use to represent the flow field. Ideally, we want  $M$  to be as large as possible to achieve maximum spatial resolution of the wind field. For example, we see from (3-10) that the smallest fluctuations resolvable in the flow field are represented in the Fourier space by wave number

$$k = \frac{2\pi M}{L_x} .$$

Consequently, if the simulation domain dimension  $L_x$  is of order 1000 km, then  $M$  must be of order 50 to resolve flow features as small as 20 km in scale. With  $M$  this large, the number of independent Fourier amplitudes in (3-20) is of the order of  $10^4$ , which is two orders of magnitude larger than the number  $J$  of meteorological stations that one would expect to find in a region of 1000 km scale. In a case such as this where the number of dependent variables exceeds the number of constraining equations, the variables are not uniquely determined. Rather, in a  $4M(M+1)$  dimensional hyperspace in which each coordinate axis is associated with one of the dependent variables in (3-20), the equation set (3-20) defines a hyperplane of dimension  $4M(M+1) - 2J$  each point of which represents a possible description of the wind field. We say that it is a possible description because it satisfies all observations exactly and it satisfies mass conservation. However, if one were to examine the flow descriptions provided by individual points on the hyperplane one would find that most

of them are improbable flows in the sense that they contain features that one does not normally observe in the atmosphere. For example, a flow field might contain an intense vortex between the wind observing sites without violating either the velocities measured at the wind stations  $(\hat{u}_j, \hat{v}_j)$  or mass conservation. Therefore, we need a means of assigning probabilities to each point on the hyperplane defined by (3-20). In particular, we need a technique for selecting only those points from the plane that describe flow fields that are likely to have occurred.

One such selection method is to search for flow fields whose kinetic energy spectra are consistent with the wind spectra observed in the atmosphere. For example, Nastrom, Gage and Jasperson (1984) found from analyses of aircraft data that in the upper troposphere both the  $u$  and  $v$  wind components have energy spectral densities that are proportional to the  $-5/3$  power of the wave number for wave numbers  $|k|$  greater than about  $5 \cdot 10^{-5} \text{ rad m}^{-1}$ , i.e., wave lengths smaller than about 500 km; and energy density proportional to  $|k|^{-3}$  for longer wave length perturbations. In the context of our flow description, these empirical spectral relationships can be expressed in the form

$$\frac{\bar{E}_{i+1}(k,l)}{\bar{E}_i(k,l)} = \left( \frac{r_{i+1}}{r_i} \right)^{-p} \quad (3-21)$$

where

$$\bar{E}_i(k,l) = \frac{1}{N_i} \sum_{k'} \sum_{l'} [|U(k',l')|^2 + |V(k',l')|^2] \quad (3-22)$$

where the summation is over all  $N_i$  wave number vectors  $(k',l')$  lying in a ring of inner and outer radii  $r_i - \Delta r/2$  and  $r_i + \Delta r/2$ , respectively, in wave number space. The power  $p$  can be either 3 or  $5/3$  depending on the magnitude of the wave number amplitude  $r$ .

We now advance the following working hypothesis:

The probability of occurrence of the atmospheric flow field described by a given point on the hyperplane defined by (3-20) is directly proportional to the degree to which the Fourier transform associated with that point satisfies (3-21) with  $p = 5/3$ . (3-23)

In order to cast this working hypothesis into an operational procedure, we must devise a method of finding those points on the hyperplane that satisfy (3-21) best. Once we have collected a sufficiently large sample of such points, say 10, we can assign each of them finite probability values which sum to 1 and assign all remaining points on the hyperplane zero probability. The finite set of probable flows then constitutes our working definition of the flow ensemble that we introduced earlier in (3-1).

The actual selection of the flow fields that will make up the ensemble is accomplished using methods of operations research. This is done by defining an optimization problem whose solutions are identical to the flow fields we seek and then applying operations research methods to solve the problem. In our case the appropriate optimization problem is the following:

Find a "design vector"  $\tilde{x}$  of dimension  $4M(M+1)$  whose elements are independent Fourier amplitudes, i.e.,

$$\tilde{x} = \begin{bmatrix} \tilde{u}_r(k,1) \\ \tilde{v}_i(k,1) \\ \tilde{u}_r(k,0) \\ \tilde{v}_i(k,0) \end{bmatrix} \quad (3-24)$$

which minimizes the objective function

$$f(\underline{x}) = \sum_{i=1}^K |f_i(\underline{x})| \quad (3-25)$$

where

$$f_i(\underline{x}) = \frac{\bar{E}_{i+1}(\underline{x})}{\bar{E}_i(\underline{x})} - \left( \frac{r_{i+1}}{r_i} \right)^p \quad (3-26)$$

and  $\bar{E}_i(\underline{x})$  is given by (3-22) and  $K$  is an integer representing the largest number of rings in  $(k,l)$  space that one can define such that each contains at least an order of 10 Fourier components  $U$  or  $V$ . In searching for the vector  $\underline{x}$  that minimizes (3-25), we confine attention to those that satisfy the following constraints:

$$(1) \quad [A]\underline{x} = \underline{b} \quad (3-27)$$

where  $[A]$  is a  $4M(M+1) \times 2J$  matrix of the coefficients  $a_j(k,l)$ ,  $b_j(k,l)$ ...,  $f_j(k)$  in Eq. 3-20 and  $\underline{b}$  is a vector of length  $2J$  whose elements are  $\hat{u}_j$ ,  $\hat{v}_j$ ;

$$(2) \quad \sum_{k=-M}^M \sum_{l=-M}^M [|U(k,l)|^2 + |V(k,l)|^2] \leq E \quad (3-28)$$

where  $E$  is a given energy value. This last constraint is added to gain control over the total kinetic energy of the flow fields that the optimization routine finds.

The optimization problem (3-25)-(3-28) is solved by first defining a secondary objective function

$$g(\underline{w}, \underline{x}) = \sum_{i=1}^{K+1} w_i \bar{E}_i(\underline{x}) = \underline{x}^T [Q] \underline{x} \quad (3-29)$$

where

$$0 < w_i < 1, \quad i=1,2,\dots,K+1 \quad (3-30a)$$

and

$$\sum_{i=1}^{K+1} w_i = 1 \quad (3-30b)$$

Eq. (3-29) can be modified to incorporate constraint (3-27) explicitly:

$$F(\underline{w}, \underline{x}, \underline{\lambda}) = \underline{x}^T [\underline{Q}] \underline{x} + \underline{\lambda}^T \{ [\underline{A}] \underline{x} - \underline{b} \} \quad (3-31)$$

where  $\underline{\lambda}$  is a vector of length  $2J$  of Lagrange multipliers. The necessary conditions for minimization of the function  $F$  are

$$\frac{\partial F}{\partial \underline{x}} = 2[\underline{Q}] \underline{x} + [\underline{A}]^T \underline{\lambda} = 0 \quad (3-32)$$

$$\frac{\partial F}{\partial \underline{\lambda}} = [\underline{A}] \underline{x} - \underline{b} = 0 \quad (3-33)$$

whose solutions are

$$\underline{\lambda} = [\underline{B}]^{-1} \underline{b} \quad (3-34)$$

$$\underline{x} = -\frac{1}{2} [\underline{Q}]^{-1} [\underline{A}]^T [\underline{B}]^{-1} \underline{b} \quad (3-35)$$

where

$$[\underline{B}] = -\frac{1}{2} [\underline{A}] [\underline{Q}]^{-1} [\underline{A}]^T. \quad (3-36)$$

The vector  $\underline{x}$  of velocity transforms given by (3-35) minimizes the secondary objective function  $g$  defined by (3-29) while satisfying the equation set (3-20) that defines the set of possible flows. However, this vector does not necessarily satisfy either constraint (3-28) or the primary objective function (3-25), which prescribes the spectral energy density. To achieve compliance with these two conditions, we vary the

vector  $\underline{w}$  using nonlinear programming procedures until (3-25) is minimized and (3-28) is satisfied. Through this combined process we arrive ultimately at a vector  $\underline{x}$  that represents a single point on the hyperplane of possible flow fields and is consistent with the empirical energy spectrum (3-21). By changing the starting value of the vector  $\underline{w}$  in (3-36) and/or the total energy  $E$  of the flow that enters in constraint (3-28), we can generate another vector  $\underline{x}$  from the flow field hyperplane that is consistent with the energy spectrum (3-21). By repeating this process we can produce an ensemble of flow fields  $\underline{x}_i$ ,  $i = 1, \dots, I$ , to each of which we can assign a probability value  $p_i$  depending on how well it satisfies the spectral condition (3-25), and scaled such that the sum of the probability values satisfies (3-2).

It is important to keep in mind that the flow ensemble produced by the process described above pertains to a single hour  $t$ , namely the hour that the observed wind velocities  $(u_j, v_j)$   $j=1, \dots, J$ , were measured. Obviously, to perform air pollution transport simulations or to compute the dispersion kernels  $P_S$  and  $P_L$ , which is the ultimate objective of our current effort, we need descriptions of the winds that span the period required for material to move from the source to the farthest receptor of interest. Thus, our next task is to extend the procedure developed here for generating a flow ensemble at a single hour to treat time intervals of arbitrary length. This is the subject of the next section. Once that step has been completed, we will have the information we need to determine the functional forms of the dispersion kernels  $P_S$  and  $P_L$ .

#### 4. Generating an ensemble of atmospheric flow fields in space-time

In the previous section we used a set of discrete wind velocity observations ( $\hat{u}, \hat{v}$ ) at  $J$  arbitrary locations in space at a single hour  $t$ , and the principle of mass conservation to define a set of possible descriptions of the continuous wind field in the space domain  $D$  at time  $t$ . Each description of the flow is represented by a  $4M(M+1)$  dimensional vector  $\underline{x}$  whose elements are Fourier amplitudes (see 3-24). Here  $M$  is an arbitrary integer determined by the number of Fourier components chosen to represent the flow field. The set of  $\underline{x}$  that constitutes possible descriptions of the wind lie on a single hyperplane defined by Eq. 3-20. We described a procedure in Eqs. 3-29 through 3-36 by which we can locate those points  $\underline{x}_i$ ,  $i=1, \dots, I$ , on the hyperplane whose kinetic energy spectra conform closely to wind energy spectra observed in the atmosphere. We assign those  $I$  flow fields finite probability values and implicitly relegate all remaining points on the hyperplane to the set of improbable flows. The  $I$  points  $\underline{x}_i$  comprise an ensemble of flows at the single time  $t$ . We want now to extend this method to produce an ensemble applicable to an arbitrary time period  $T$ .

Our approach consists of two steps. First, we apply the techniques developed in the previous section, and summarized above, to obtain an ensemble of  $I$  flow fields at each of the hours in the period  $T$  for which data are available. For example, suppose that  $T$  is  $N$  hours long and that wind data are available at hourly intervals. In this case we can create  $N+1$  ensembles

$$\underline{x}_i^n, \quad i=1, \dots, I; \quad n=0, 1, \dots, N \quad (4-1)$$

where  $\tilde{x}_i^n$  is a collection of flow descriptions for hour  $t_n$ . (As we noted earlier both the number  $J$  and the locations of the wind observations may change from one hour to another.) The second step in creating the space-time ensemble in  $(D,T)$  is to find those sequences of flows in the set (4-1), i.e.,

$$[\tilde{x}_\alpha^0, \tilde{x}_\beta^1, \tilde{x}_\gamma^2, \dots, \tilde{x}_\phi^N]_k, \quad 1 \leq \alpha, \beta, \gamma, \dots, \phi \leq I; \quad (4-2)$$

$$1 \leq k \leq K,$$

that are most consistent with the momentum equation. Here  $K$  is an integer, generally of the same order as  $I$ , which represents the total number of flow fields that we want in the space-time ensemble. The upper bound on  $K$  is the number of unique sequences of the form (4-2) that can be constructed from the  $(N+1)$  subensembles  $\tilde{x}_i^N$ . In the present case we have

$$K \leq K_{\max} = I^{N+1}$$

With  $I$  of order 10 and  $N$  of order 24, we see that  $K_{\max}$  is an astronomically large number.

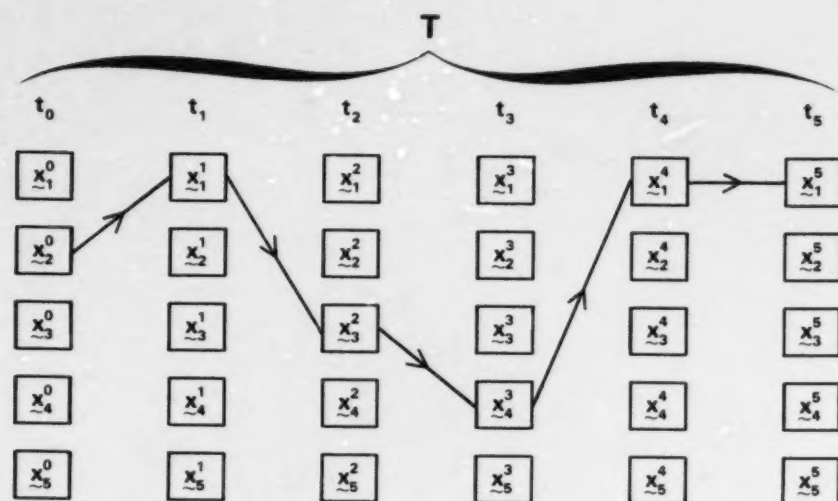


Figure 2. Schematic representation of a sequence of flow fields that constitutes a single member of an ensemble of flows in space time  $(D, T)$ .

The two-step process for generating a space-time ensemble of flow fields is illustrated in Figure 2 for the case where  $I=N=5$ . Line segments connecting members of the subensembles  $x_i^n$  in the figure represent the sequence  $[x_2^0, x_1^1, x_3^2, x_4^3, x_1^4, x_1^5]$ , which is a possible member of the space-time ensemble. In this particular flow the velocity field evolves from state  $x_2^0$  at time  $t_0$  to state  $x_1^1$  at the next hour  $t_1$ , to state  $x_3^2$  at hour  $t_2$  and so on, reaching state  $x_1^5$  at the final hour  $t_5$ . Each of the states  $x_i^n$  illustrated in the figure is defined in terms of the wind observations at hour  $t_n$ , the mass conservation principle, and an empirical energy spectrum relationship. Indeed, the set of states  $x_i^n, i=1, \dots, I$ , at each hour  $t_n$  represents what we regard to be the most probable descriptions of the flow at that hour. What we require now is a means of determining the most probable sequences of these states. Toward

this end we advance the following complement to hypothesis (3-23):

The probability of a given sequence of flow fields  $[x_{\alpha}^0, x_{\beta}^1, \dots, x_{\psi}^N]$  is directly proportional to the degree to which that sequence satisfies the principle of momentum conservation. (4-3)

To implement this hypothesis we will use dynamic programming and Bellman's principle of optimality.

This procedure is described best using an example. Consider any one of the flow states  $x_i^n$ . According to the momentum conservation principle, this state will evolve into a new state, say  $y_i^{n+1}$ , at time  $t_{n+1}$  where

$$y_i^{n+1} = F(x_i^n, \xi) \quad (4-4)$$

and  $\xi$  is a vector of specified independent parameters such as boundary conditions, forcing fields, etc, and the function  $F$  represents the solution of the momentum equation. We now define what we will call a penalty function  $\epsilon_{ij}^n$ :

$$\epsilon_{ij}^n = ||x_i^{n+1} - x_j^{n+1}||, \quad i < i, j < I \quad (4-5)$$

where  $||$  denotes a suitable norm, such as the sum of the squares of the vector components. In words, the penalty function  $\epsilon_{ij}^n$  is a measure of the difference between the flow state that  $x_i^n$  would in theory become at time  $t_{n+1}$  and the  $j$ -th member  $x_j^{n+1}$ ,  $1 < j < I$ , of our ensemble of states at  $t_{n+1}$ . Ideally,  $\epsilon_{ij}^n$  should be zero for one value of  $j$  but in general this will never occur due to errors in the wind data and in the specifications of  $\xi$ , etc.

Let

$$\epsilon_i^{n*} = \min_j \{\epsilon_{ij}^n\} . \quad (4-6)$$

It follows that if  $x_i^n$  has been selected as the flow field description at time  $t_n$ , the best description at time  $t_{n+1}$  is  $x_{\alpha}^{n+1}$ , where  $\alpha$  is the state for which the penalty function is minimum, namely

$$\epsilon_{i\alpha}^n = \epsilon_i^{n*} . \quad (4-7)$$

The dynamic programming procedure starts at the next to the last time step  $t_{N-1}$  and computes the minimum penalty functions for each of the  $I$  members of the subensemble  $x_i^{N-1}$ , i.e.,

$$\epsilon_i^{(N-1)*} , \quad 1 \leq i \leq I. \quad (4-8)$$

These values tell us which state at time  $t_N$  is most consistent, from the standpoint of the momentum equation, with a given state  $x_i^{N-1}$  at time  $t_{N-1}$ . Once the minimum values (4-8) have been calculated, we move backward one step to time  $t_{N-2}$ . At that point the penalty function is defined by (cf 4-5)

$$\epsilon_{ij}^{N-2} = ||x_i^{N-1} - x_j^{N-1}|| + \epsilon_j^{(N-1)*} \quad (4-9)$$

In words, the penalty function associated with a given pair of states  $x_i^n$  and  $x_j^{n+1}$  separated in time by a single time step is the minimum penalty that one could achieve by beginning at state  $i$  at time  $n$  and proceeding through state  $j$  at time  $n+1$  all the way to the final time level  $t_N$ . In analogy with (4-6) the minimum penalty for a given state  $x_i^{N-2}$  is

$$\epsilon_i^{(N-2)*} = \min_j \{\epsilon_{ij}^{N-2}\} \quad (4-10)$$

The definitions of the penalty function (4-9) and minimum penalty (4-10) can be applied at time step  $N-3$  and to each earlier time level in turn until we arrive finally at values for the initial hour, namely

$$\epsilon_{ij}^0 \text{ and } \epsilon_i^{(0)*} . \quad (4-11)$$

Taking the minimum of the latter over all  $I$ , i.e.,

$$\min_i \{ \epsilon_i^{(0)*} \} \quad (4-12)$$

we obtain the information necessary to define the optimal sequence of states  $\underline{x}_i^n$  in space-time. Using the example illustrated in Figure 2, suppose that  $i=2$  is the state for which  $\epsilon_i^{(0)*}$  is a minimum. Now by definition  $\epsilon_i^{(0)*}$  is the minimum of all the penalty values  $\epsilon_{ij}^{(0)}$ ,  $1 \leq j \leq I (=5)$  that are associated with state  $i$ . Suppose that  $j=1$  is the state for which  $\epsilon_{ij}^{(0)} = \epsilon_i^{(0)*}$ . In this case the pair of states  $\underline{x}_2^0, \underline{x}_1^1$  are the two that are most consistent with the momentum equation. Looking next at state  $i=1$  at time  $t_1$ , we know the associated minimum penalty  $\epsilon_i^{(1)*}$  and we know with which state  $j$  at time  $t_2$  it is associated --  $\epsilon_i^{(1)*} = \epsilon_{13}^{(1)}$  in the present example. This information tells us that the sequence  $[\underline{x}_2^0, \underline{x}_1^1, \underline{x}_3^2]$  is optimal from the momentum conservation standpoint during the first three time steps. Proceeding in this way we arrive finally at the full sequence illustrated by the connecting line segments in Figure 2.

The beauty of dynamic programming in this particular problem is that it is able to determine the optimal sequence by evaluating only  $N \cdot I^2$  of the possible sequences. By contrast a straightforward approach would have to examine all  $K_{\max}$  sequences, which total

$$K_{\max} = I^{N+1} .$$

For example, in the case where  $I=10$  and  $N=24$ , dynamic programming arrives at the optimal sequence after testing only 2400 of the  $10^{25}$  possibilities! The final step in the implementation of the dynamic programming procedure is to express the state vectors  $y_1^n$ , which enter in the definition of the penalty function (see 4-9), in terms of the momentum equation.

Since the state vectors  $x_1^n$  are expressed in terms of complex Fourier amplitudes, we must express the solutions  $y_1^n$  of the momentum equation in the same form. We start with the equations for a 2-D fluid in physical coordinates:

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} = - \frac{1}{\rho} \frac{\partial p}{\partial x} + \kappa \left( \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) \quad (4-13a)$$

$$\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} = - \frac{1}{\rho} \frac{\partial p}{\partial y} + \kappa \left( \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} \right) \quad (4-13b)$$

where  $p$  is pressure,  $\rho$  is the fluid density and  $\kappa$  is the kinematic viscosity. Eqs. 4-13 neglect frictional and external forces on the fluid.

It is convenient to work with the vorticity

$$\Gamma = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \quad (4-14)$$

whose governing equation is found from (4-13) and the mass continuity equation (3-14) (with  $\delta = 0$ ) to be

$$\frac{\partial \Gamma}{\partial t} + u \frac{\partial \Gamma}{\partial x} + v \frac{\partial \Gamma}{\partial y} = \kappa \left( \frac{\partial^2 \Gamma}{\partial x^2} + \frac{\partial^2 \Gamma}{\partial y^2} \right) \quad (4-15)$$

In terms of the complex Fourier amplitudes ( $U, V$ ) of the velocity ( $u, v$ ),

the vorticity is expressed in the form (see 3-8 and 3-9)

$$\Gamma(x,y,t) = i \sum_{k=-M}^M \sum_{l=-M}^M [k_{\mu} V(k,l,t) - l_{\nu} U(k,l,t)] e^{ip} \quad (4-16)$$

where

$$p = k_{\mu} x + l_{\nu} y \quad (4-17)$$

Substituting (4-16) into (4-15) and making use of (3-17) we obtain after a considerable amount of algebra the equations governing the temporal behavior of U and V:

$$\begin{aligned} -i \frac{(k_{\mu})^2 + (l_{\nu})^2}{l_{\nu}} \frac{\partial U(k,l,t)}{\partial t} + A(k,l,t) + B(k,l,t) \\ - i \kappa \frac{[(k_{\mu})^2 + (l_{\nu})^2]^2}{l_{\nu}} U(k,l,t) = 0 \end{aligned} \quad (4-18a)$$

$$l \neq 0;$$

$$\begin{aligned} i k_{\mu} \frac{\partial V(k,o,t)}{\partial t} + A(k,o,t) + B(k,o,t) + i \kappa (k_{\mu})^3 V(k,o,t) = 0 \\ l = 0. \end{aligned} \quad (4-18b)$$

In Eq. (4-18)

$$A(k,l,t) = \sum_{\substack{k''= \\ \max(-M, -M+k)}}^{\min(M+k, M)} \sum_{\substack{l''= \\ \max(-M, -M+l)}}^{\min(M+l, M)} U(k'',l'',t) G_1(k-k'', l-l'', t) \quad (4-19a)$$

$$G_1(k',l',t) = - [V(k',l',t) k'_{\mu} - U(k',l',t) l'_{\nu}] k'_{\mu} \quad (4-19b)$$

and

$$B(k,l,t) = \sum_{\substack{k''= \\ \max(-M, -M+k)}}^{\min(M+k, M)} \sum_{\substack{l''= \\ \max(-M, -M+l)}}^{\min(M+l, M)} V(k'',l'',t) G_2(k-k'', l-l'', t) \quad (4-20a)$$

$$G_2(k', l', t) = - [V(k', l', t)k'_\mu - U(k', l', t)l'_\nu]l'_\nu \quad (4-20b)$$

If we initialize Eq. (4-18) with the values of the Fourier amplitudes at time  $t_n$ , i.e., we use as initial condition (cf 3-24)

$$\underline{x}^n = \begin{bmatrix} \underline{u}_r(k, l, t_n) \\ \vdots \\ \underline{v}_i(k, 0, t_n) \end{bmatrix} \quad (4-21)$$

we can integrate (4-18) forward in time to obtain

$$\underline{y}^{n+1} = \begin{bmatrix} \underline{u}_r(k, l, t_{n+1})_p \\ \vdots \\ \underline{v}_i(k, 0, t_{n+1})_p \end{bmatrix} \quad (4-22)$$

where the subscript "p" in (4-22) signifies that the Fourier components are predicted by the momentum equation rather than deduced from the diagnostic data, which was the method we used in Section 3 to compile the ensemble of states  $\underline{x}_i^n$ . It is the predicted state vector  $\underline{y}^{n+1}$  obtained from (4-18) for each of the states  $\underline{x}_i^n$  that we require in (4-5) to compute the penalty functions. Thus, all of the information that we need to apply the dynamic programming procedure is now in place. The final result is an ensemble of K velocity fields in space-time, namely sequences of flow states of the form

$$[\underline{x}^0, \underline{x}^1, \dots, \underline{x}^N]_k \quad 1 \leq k \leq K \quad (4-23)$$

where  $\underline{x}^n$  is one of the  $4M(M+1)$  dimensional vectors of Fourier amplitudes derived in Section 3 that provides a continuous description of the flow in space at time  $t_n$ ; and where each of the  $\underline{x}^n$  may vary from one member  $k$  of the ensemble to another. The ensemble of flows represented by (4-23) is the set required to compute the dispersion kernel functions  $P_S$  and  $P_L$  (see Section 2, Eq. 2-14, 15). In the next section we begin the formulation of the optimization technique that will use the kernels  $P_S$  and  $P_L$  to find optimal sites for pollutant sources.

## 5. Optimal solutions to the plant siting problem

In the last Sections we focused on the problem of formulating the kernel functions that describe the impact of pollutant emissions on air quality. In this section we use those functions and additional information to find locations for power plants in a given region that achieve an optimal trade-off between cost and air quality deterioration. As we noted earlier, for academic interest we will consider here the case where minimizing pollutant concentrations is one of several objectives, rather than a constraint, on the design of the overall system. The latter would be the proper manner of treatment under current pollutant control regulations. The procedure that we develop here to solve the optimization problem can handle the air quality aspects of the problem as either a constraint or an objective.

Let us first define the variables that we will use in formulating the problem mathematically:

$\tilde{z}$  = vector of design variables

$$= \begin{bmatrix} \tilde{z}_L \\ \hline \tilde{z}_F \end{bmatrix} \quad (5-1)$$

where

$\tilde{z}_F$  = vector of fuel allocation parameters

$\tilde{z}_L$  = vector of coordinates of the power plant sites;

$m$  = number of sources of fuel, such as coal mines;

$n$  = number of power plants to be located;

$x_{ij}$  = flow rate of fuel from mine  $i$  to plant  $j$ ;

$d_{ij}$  = upper bound restriction on the flow rate  $x_{ij}$ ;

$C_{ij}$  = unit cost of fuel at mine  $i$  plus transportation cost to plant  $j$ ;

$m_i$  = maximum output capacity of fuel at mine  $i$ ;

$r_j$  = thermal energy requirement of plant  $j$ ;

$k_i$  = calorific value of fuel from mine  $i$ ;

$s_i$  = percentage sulfur in fuel from mine  $i$ ;

$a_i$  = percentage ash content in fuel from mine  $i$ ;

$(\bar{x}_i, \bar{y}_i)$  = location of mine  $i$ ;

$(x_j, y_j)$  = location of plant  $j$ .

The optimal plant sites must be determined subject to the following constraints:

- (1) the thermal energy requirements of each plant must be met, i.e.,

$$\sum_{i=1}^m k_i x_{ij} > r_j, \quad j=1,2,\dots,n; \quad (5-2)$$

- (2) the flow rate of fuel between any mine and any plant must not exceed the limiting rate of fuel transportation, i.e.,

$$0 < x_{ij} < d_{ij}, \quad \text{all } i, j; \quad (5-3)$$

- (3) the total rate of consumption of fuel from any mine cannot exceed the mines maximum output capacity, i.e.,

$$\sum_{j=1}^n x_{ij} < m_i, \quad i=1,2,\dots,m; \quad (5-4)$$

- (4) plant sites must fall within the designated feasible areas, i.e.,

$$(x_j, y_j) \in R_j, \quad j=1,2,\dots,n \quad (5-5)$$

Each of the above constraints can be expressed in the general form

$$G_l(\underline{z}) < 0, \quad l=1,\dots,M \quad (5-6)$$

where  $M$  is the total number of constraints.

Our search is for locations  $(x_j, y_j)$  of the  $n$  power plants that achieve optimal satisfaction of five separate objectives. These are

Objective (1): Minimization of fuel cost for all plants combined.

The purchase price plus transportation charges of fuel for plant  $j$  expressed as cost per unit time is

$$C_j = \sum_{i=1}^m C_{ij} X_{ij}. \quad (5-7)$$

Our objective is to minimize the total cost of all plants, namely minimize

$$F_1(z) = \sum_{j=1}^n C_j = \sum_{j=1}^n \sum_{i=1}^m C_{ij} X_{ij} \quad (5-8)$$

Here the cost factor  $C_{ij}$  can be expressed as

$$C_{ij} = C_i^f + C_{ij}^t D_{ij} \quad (5-9)$$

where

$C_i^f$  = cost per unit mass of fuel at mine  $i$ ,

$C_{ij}^t$  = cost of transportation of fuel per unit mass per unit distance along the line connecting mine  $i$  and plant  $j$ ,

$D_{ij}$  = transport distance which, for the purposes of this study, we will take to be

$$= [(\bar{x}_i - x_j)^2 + (\bar{y}_i - y_j)^2]^{1/2} \quad (5-10)$$

Objective (2): Minimization of short-term sulfur dioxide concentration.

Using expression (2-14) we can express this objective in the form

$$\text{minimize } F_2(\underline{z}) = \left[ \sum_{j=1}^n \sum_{i=1}^m x_{ij} s_i \right] \max_{t \in T_L} \{P_s(t; x_{s1}, \dots, x_{sn})\} \quad (5-11)$$

where the fuel flow rates  $x_{ij}$  are expressed in units of mass/time.

Objective (3): Minimization of long-term sulfur dioxide concentration.

The long-term averaged concentration of a pollutant species is given by (2-15). Using that equation we can write the present objective in the form

$$\text{minimize } F_3(\underline{z}) = \left[ \sum_{j=1}^n \sum_{i=1}^m x_{ij} s_i \right] P_L(x_{s1}, \dots, x_{sn}) \quad (5-12)$$

The remaining two objectives are similar to (2) and (3) except they apply to minimization of particulate concentration, namely

Objective (4): Minimize short-term averaged particulate concentration, i.e.,

$$\text{minimize } F_4(\underline{z}) = Q_a \max_{t \in T_L} \{P_s(t; x_{s1}, \dots, x_{sn})\} \quad (5-13)$$

Objective (5): Minimize long-term averaged particulate concentration, i.e.,

$$\text{minimize } F_5(\underline{z}) = Q_a P_L(x_{s1}, \dots, x_{sn}). \quad (5-14)$$

In (5-13) and (5-14)

$$Q_a = \sum_{j=1}^n \sum_{i=1}^m f_j (1 - E_j) x_{ij} a_i \quad (5-15)$$

where  $f_j$  is the fraction of fly ash that leaves the furnace at plant  $j$  and  $E_j$  is the efficiency of the particulate collecting equipment installed at plant  $j$ .

Our aim is to develop a procedure that will allow us to achieve optimal satisfaction of all five objectives (5-8, 11, 12, 13, and 14) jointly while maintaining compliance with the system constraints (5-2, 3, 4, and 5). This

is termed a multi-criteria optimization problem, and we shall attempt now to solve it using concepts borrowed from game theory.

Game theory has been developed for both cooperative and noncooperative games. In the latter, each player acts independently in an effort to maximize his own gain. The outcome of such games is known as the Nash equilibrium solution to the nonzero-sum game. In the special case where only two players are involved and one's gain is the other's loss, the game is called a zero sum game and the associated Nash equilibrium solution reduces to the so-called saddle-point solution.

In cooperative games, players form coalitions under the expectation that working together each can achieve an outcome that is better than the Nash solution, i.e., the gain that each would have realized had they competed individually to maximize their gains. The measure of success of cooperative play is embodied in the concept of pareto-optimality which involves a single objective function defined by

$$F(\underline{z}) = \sum_{i=1}^N w_i F_i(\underline{z}) \quad (5-16a)$$

where  $F_i$  are objective functions associated with each of the  $N$  players and  $w_i$  is a weighting function with the properties

$$\begin{aligned} w_i &> 0, \quad i=1, \dots, N \\ \sum_{i=1}^N w_i &= 1. \end{aligned} \quad (5-16b)$$

Given a weighting function  $w_i$ ,  $i=1, \dots, N$ , a solution  $\underline{z}$  that minimizes  $F(\underline{z})$ , as defined by (5-16a), is called a pareto-optimal solution. However, this solution is clearly a function of the weighting function  $w$ ; and since some

w give more weight to the objective function of some players than others, implementation of the pareto-optimal objective requires that a function  $w_i$ ,  $i=1, \dots, N$ , be found that all players will agree yields a fair overall objective  $F(z)$  of the game.

To find such a weighting function we begin by considering the situation where each of the  $N$  players is allowed in turn to find the design vector  $\underline{z}$  that minimizes his objective function alone subject to the system constraints. Thus, we define the vector

$$\underline{z}_i^* = \text{value of } \underline{z} \text{ for which } F_i(\underline{z}) = \text{minimum.}$$

By definition  $F_i(\underline{z}_i^*)$  is a measure of the best outcome that player  $i$  could ever achieve in the game, but it is generally not an optimal outcome for any of the other players. That is,

$$F_j(\underline{z}_i^*) \neq \text{minimum for } i \neq j.$$

Using the  $\underline{z}_i^*$ ,  $i=1, \dots, N$ , and the objective functions  $F_i(\underline{z})$ ,  $i=1, \dots, N$ , evaluated at these points we can form the matrix

$$[M] = \begin{bmatrix} F_1(\underline{z}_1^*) & F_2(\underline{z}_1^*) & \dots & F_N(\underline{z}_1^*) \\ F_1(\underline{z}_2^*) & F_2(\underline{z}_2^*) & \dots & F_N(\underline{z}_2^*) \\ \vdots & \vdots & \ddots & \vdots \\ F_1(\underline{z}_N^*) & F_2(\underline{z}_N^*) & \dots & F_N(\underline{z}_N^*) \end{bmatrix} \quad (5-18)$$

This matrix contains useful information in that the smallest element in column  $i$  is a measure of the best that player  $i$  can hope to achieve in the game under any conditions while the largest element in that column is the worst that player  $i$  could do in cooperative play. Consequently, we are led

to define the "super-criterion"

$$S(\underline{w}) = \prod_{i=1}^N \frac{\bar{F}_i}{F_i(\underline{z}_C)} \quad (5-19)$$

where  $\bar{F}_i$  is the largest element in column  $i$  in  $[M]$  and  $\underline{z}_C$  is the design vector that minimizes

$$F_C(\underline{z}) = \sum_{i=1}^N w_i \frac{F_i(\underline{z})}{F_i^*} \quad (5-20)$$

where  $F_i^*$  is the smallest element in column  $i$  in  $[M]$ . Note that  $S(\underline{w})$  is a measure of how far each player is from his worst outcome when the design vector is selected as the pareto-optimal solution based on the weighting function  $\underline{w} = (w_1, \dots, w_N)$ . In other words, in our search for a weighting function  $\underline{w}$  that all players will agree best combines the individual objectives into a single objective  $F$ , given by (5-16), we are led to the function  $\underline{w}$  that maximizes the super criterion (5-19). Because with that particular  $\underline{w}$  player  $i$  is assured that in the specification of a single function  $F$  that best represents the objective of the coalition of players that he has joined, his individual objective  $F_i$  is given a weight  $w_i$  that will maximize his position in the game relative to the worst outcome he could suffer as a result of cooperative play.

Implementation of the pareto-optimal procedure is performed in the following steps:

- Step 1. Construct the matrix  $[M]$ , (5-18), by solving  $N$  single criterion optimization problems.

- Step 2. Choose any weighting function  $w$  that satisfies (5-17) and use it in (5-20) to find the design vector  $\underline{z}_c$  for which  $F_c(\underline{z})$  is a minimum.
- Step 3. Use the vector  $\underline{z}_c$  obtained in Step 2 to evaluate the super-criterion (5-19). Determine the change in  $\underline{z}_c$  that would result in the largest increase in the super criterion.
- Step 4. Iterate between Steps 2 and 3 until the super criterion has reached a maximum value.

In the next Section we will apply this procedure to a sample problem.

6. Application of the optimization procedures to a hypothetical problem.

In this section we apply the techniques that we developed in each of the previous sections to a rather elementary problem involving the siting of two power plants in a region containing two sources of coal. The objective is to find plant sites that minimize operating costs and air quality impacts jointly given information on the meteorology of the region, the cost and pollutant contents of the coal from each mine, etc. As we noted in the introduction, the principal aim of this report is technique development rather than application, and therefore the problem that we will consider here is the simplest that will allow us to demonstrate key elements of the optimization procedure.

We divide the problem into four parts along the same lines that we used in the earlier sections to address the optimization procedure. In the first subsection we consider the creation of the subensembles of flows  $\tilde{x}_1^n$  at each hour of the simulation period, followed in subsection 6.2 with an application of the dynamic programming technique to form a space-time flow ensemble. In subsection 6.3 we compute the associated dispersion kernels for the short and long-term concentration averages, and finally in 6.4 we combine the dispersion formulas and the optimization method developed in Section 5 to find optimal sites for the two hypothetical power plants.

6.1 Generation of subensembles of flow fields  $\underline{x}_i^n$  for each time interval in the simulation period  $T_L$

Let us assume that within the spatial region of interest wind data are available from seven sites. The locations of these sites and the wind speed and direction measurements made at each of them at regular intervals during the simulation period  $T_L$  are given in Table 1. The data contained in the table are actual surface wind observations made at hourly intervals at stations in the Northeastern United States during a summer month. One of the common problems with data, which is evident in Table 1, is that for one reason or another, values are frequently "missing". The technique that we developed in Section 3 for treating the winds is unaffected by either the loss or gain of data points from one observation time to another.

Table 1. Wind speed and direction observations at hourly intervals at seven surface meteorological stations. A dash (-) in a data column indicates that the value was not reported.

| Station | Coordinates | Wind Speed (m/sec)/wind direction (deg) |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|---------|-------------|-----------------------------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|         |             | (x)                                     | (y) | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  |     |     |     |     |     |     |     |     |
| BDS     | 71.03 42.37 | 6.2                                     | 190 | -   | -   | 6.2 | 190 | 5.7 | 200 | 5.7 | 200 | 5.7 | 200 | 5.1 | 190 | 5.1 | 200 | 5.1 | 230 |     |     |
| EWB     | 70.97 41.68 | 4.6                                     | 190 | 5.1 | 190 | -   | 6.2 | 190 | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   |     |     |
| FWH     | 70.52 41.65 | 7.2                                     | 210 | 7.2 | 210 | -   | 5.1 | 220 | 5.1 | 200 | 6.2 | 210 | 7.2 | 220 | 9.3 | 210 | 8.2 | 210 | 9.3 | 240 |     |
| HVA     | 70.28 41.67 | 5.1                                     | 190 | 3.6 | 230 | 5.1 | 230 | 5.1 | 220 | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   |     |     |
| NZW     | 70.93 42.15 | 2.6                                     | 180 | 3.1 | 180 | 4.1 | 180 | 4.1 | 180 | 4.1 | 180 | 4.1 | 190 | 3.6 | 180 | 3.1 | 180 | 3.1 | 170 | 2.1 | 180 |
| OWD     | 71.18 42.18 | 5.1                                     | 200 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   |     |
| PVD     | 71.43 41.73 | 6.7                                     | 210 | 6.2 | 190 | 5.7 | 210 | -   | -   | -   | -   | -   | -   | 3.6 | 190 | -   | -   | 4.6 | 210 | -   | -   |

The first step is to construct subensembles  $\underline{x}_1^n$  (see 3-24) of the wind field for each of the hours  $n=1,2,\dots,10$ . We will represent the flows in Fourier space using  $M = 8$  wave numbers (see 3-8,9) and we will generate  $I = 4$  members of each subensemble. Each member  $\underline{x}_i^n$   $i = 1,\dots,4$ ,  $n = 1,\dots,10$ , must have the following properties: (1) it satisfies the observations given in Table 1 at hour  $n$ ; (2) it satisfies the mass conservation equation (3-14), with  $\delta = 0$ ; (3) it has an energy spectrum of the form (3-21) with  $p = 5/3$ ; and (4) the total kinetic energy of the flow  $E_i$ ,  $i = 1,\dots,4$ , is within a given bound.

Recalling Section 3, we point out that a flow field  $\underline{x}_1^n$  satisfying all the conditions just stated is derived from Eq. 3-35 in which  $[Q]$  is a function of the vector  $\underline{w}$  defined implicitly in (3-29), and where  $[A]$ ,  $[B]$  and  $\underline{b}$  are determined by the wind observations and the coordinates of the measurement sites. In the present example we generate four members  $i = 1,\dots,4$  of each subensemble by varying the bound on the total kinetic energy  $E$  of the flow. It turns out that an energy spectrum of the form

$$E(k) \sim |k|^{-5/3} \quad (6-1)$$

apparently does not exist for arbitrary values of the total energy  $E$ , at least in the present limited Fourier representation of the flow (i.e., with  $M = 8$ ). This can be seen in Table 2 where we show the value of the objective function  $f(\underline{x})$ , (3-25), and the total kinetic energy  $E_i$  of each realization  $\underline{x}_i^n$ ,  $i = 1,\dots,4$ , in the subensemble at each hour  $n = 1,\dots,10$ . By definition  $f(\underline{x})$  would be zero if the energy spectrum had the form (6-1). At each of the ten hours the realization labeled "1" in Table 2 is the flow field with least kinetic energy that satisfies the continuity equation

and the wind observations. In other words,  $\tilde{x}_1^n$  is the point on the hyperplane of solutions that is closest to the origin of the phase space of Fourier amplitudes  $[\underline{u}(k), \underline{v}(k)]$ . However, this particular flow field does not conform well to the energy spectrum relationship (6-1) that we expect based on wind observations. Under our earlier hypothesis (3-23) we would assign the highest probability  $p$  to member 2 of each subensemble, since it is that member that conforms best to the spectrum (6-1); and lower probabilities to each of the other three members. However, our concern here is not with the flow field at a single hour but with the flow over the entire 10 hour simulation period  $T_L$ . To create an ensemble of flows in space-time from the set of 10 subensembles that we just generated, we use the momentum equation and dynamic programming procedure developed in Section 4. This is our next task.

Table 2. Components of the vector  $\underline{w}$  that define the matrix  $[Q]$  which in turn defines each member  $x^n$  of the subensemble of flows at each hour  $n$  through Eqs. 3-29 and 3-35. Shown also is the value of the objective function  $f(\underline{x})$  defined by (3-25) which is a measure of how well the flow field satisfies the energy spectrum (6-1). The total kinetic energy  $E_i$  of the flow is given in column four.

| Realization or<br>member number $i$ | Vector $\underline{w}$ |        |        |        | $f(\underline{x})$<br>Eq. 3-25 | Total kinetic<br>energy $E_i$ |
|-------------------------------------|------------------------|--------|--------|--------|--------------------------------|-------------------------------|
|                                     | $w_1$                  | $w_2$  | $w_3$  | $w_4$  |                                |                               |
| Hour $n = 1$                        |                        |        |        |        |                                |                               |
| 1                                   | 1.872                  | 7.902  | 9.537  | 11.243 | 3.03                           | 0.612                         |
| 2                                   | 1.267                  | 7.901  | 9.536  | 11.242 | 0.005                          | 0.708                         |
| 3                                   | 1.269                  | 8.406  | 10.614 | 12.670 | 0.92                           | 0.740                         |
| 4                                   | 1.248                  | 7.902  | 14.0   | 16.865 | 2.86                           | 0.797                         |
| Hour = 2                            |                        |        |        |        |                                |                               |
| 1                                   | 1.536                  | 5.82   | 6.578  | 6.472  | 3.09                           | 1.33                          |
| 2                                   | 1.023                  | 5.8206 | 6.5779 | 6.4719 | 0.0093                         | 1.504                         |
| 3                                   | 0.915                  | 5.795  | 6.571  | 6.472  | 1.24                           | 1.579                         |
| 4                                   | 1.024                  | 5.82   | 9.867  | 10.0   | 3.01                           | 1.663                         |
| Hour = 3                            |                        |        |        |        |                                |                               |
| 1                                   | 1.919                  | 6.829  | 9.142  | 9.172  | 3.09                           | 1.043                         |
| 2                                   | 1.279                  | 6.829  | 9.142  | 9.172  | 0.014                          | 1.177                         |
| 3                                   | 1.2748                 | 7.167  | 10.504 | 10.881 | 1.1                            | 1.236                         |
| 4                                   | 1.279                  | 6.829  | 13.713 | 16.0   | 3.78                           | 1.341                         |
| Hour = 4                            |                        |        |        |        |                                |                               |
| 1                                   | 2.495                  | 8.058  | 11.109 | 10.706 | 3.15                           | 1.018                         |
| 2                                   | 1.663                  | 8.058  | 11.109 | 10.706 | 0.014                          | 1.137                         |
| 3                                   | 1.0                    | 5.0    | 7.7    | 8.0    | 1.23                           | 1.194                         |
| 4                                   | 1.663                  | 8.058  | 16.664 | 18.0   | 3.39                           | 1.29                          |
| Hour = 5                            |                        |        |        |        |                                |                               |
| 1                                   | 1.566                  | 5.627  | 9.208  | 10.878 | 3.21                           | 0.057                         |
| 2                                   | 1.043                  | 5.626  | 9.208  | 10.878 | 0.018                          | 0.067                         |
| 3                                   | 1.052                  | 5.777  | 10.349 | 12.237 | 0.74                           | 0.07                          |
| 4                                   | 1.044                  | 5.627  | 12.0   | 16.317 | 2.52                           | 0.0769                        |

Table 2 (cont'd)

| Realization or<br>member number i | Vector $\tilde{w}$ |        |         |         | $f(x)$<br>Eq. 3-25 | Total kinetic<br>energy $E_i$ |
|-----------------------------------|--------------------|--------|---------|---------|--------------------|-------------------------------|
|                                   | $w_1$              | $w_2$  | $w_3$   | $w_4$   |                    |                               |
| Hour = 6                          |                    |        |         |         |                    |                               |
| 1                                 | 2.221              | 8.91   | 12.25   | 14.9    | 3.2                | 0.107                         |
| 2                                 | 1.4806             | 8.9085 | 12.2451 | 14.9003 | 0.005              | 0.126                         |
| 3                                 | 1.4359             | 8.6448 | 13.1148 | 17.412  | 0.93               | 0.132                         |
| 4                                 | 1.481              | 8.91   | 18.368  | 20.0    | 3.41               | 0.1428                        |
| Hour = 7                          |                    |        |         |         |                    |                               |
| 1                                 | 1.437              | 5.659  | 8.648   | 10.265  | 3.2                | 0.414                         |
| 2                                 | 0.9583             | 5.6586 | 8.6477  | 10.265  | 0.0089             | 0.487                         |
| 3                                 | 0.9115             | 5.3865 | 8.7323  | 12.0395 | 1.03               | 0.512                         |
| 4                                 | 0.958              | 5.659  | 12.972  | 14.0    | 3.41               | 0.564                         |
| Hour = 8                          |                    |        |         |         |                    |                               |
| 1                                 | 2.173              | 7.747  | 9.587   | 11.12   | 3.25               | 0.707                         |
| 2                                 | 1.4493             | 7.747  | 9.5866  | 11.1113 | 0.013              | 0.802                         |
| 3                                 | 1.0482             | 5.9949 | 7.5996  | 9.7982  | 1.29               | 0.842                         |
| 4                                 | 1.449              | 7.747  | 14.38   | 16.0    | 3.29               | 0.892                         |
| Hour = 9                          |                    |        |         |         |                    |                               |
| 1                                 | 1.778              | 6.379  | 8.328   | 9.624   | 3.2                | 0.562                         |
| 2                                 | 1.1854             | 6.3791 | 8.3275  | 9.6242  | 0.01               | 0.641                         |
| 3                                 | 1.1184             | 6.044  | 9.2284  | 10.9533 | 1.02               | 0.674                         |
| 4                                 | 1.185              | 6.379  | 12.491  | 14.0    | 3.29               | 0.72                          |
| Hour = 10                         |                    |        |         |         |                    |                               |
| 1                                 | 1.606              | 6.088  | 9.482   | 12.492  | 3.15               | 0.849                         |
| 2                                 | 1.0707             | 6.0877 | 9.4822  | 12.4915 | 0.07               | 0.997                         |
| 3                                 | 0.9988             | 5.6662 | 9.4749  | 13.7029 | 0.87               | 1.046                         |
| 4                                 | 1.071              | 6.088  | 9.482   | 18.737  | 2.52               | 1.08                          |

## 6.2 Generation of an ensemble of flows in space-time using the momentum equation and dynamic programming

Using each of the  $\tilde{x}_j^n$  given in Table 2 we computed the penalty function  $\epsilon_{ij}^n$  defined in the manner of (4-9) for each pair of states  $(\tilde{x}_i^n, \tilde{x}_j^{n+1})$ . In these calculations the predicted states  $\tilde{x}_i^n$  were obtained from the vorticity equations (4-18a, b) using a simple forward difference for the time derivatives, i.e.,

$$\frac{\partial U(k,l,t_n)}{\partial t} = \frac{[U(k,l,t_{n+1}) - U(k,l,t_n)]}{(t_{n+1} - t_n)} \quad (6-2)$$

After the penalty functions were computed for each time level  $n = 1, \dots, 10$ , dynamic programming was applied to arrive at the following, most probable space-time sequence of flow:

$$\tilde{x}_1^1, \tilde{x}_1^2, \tilde{x}_1^3, \tilde{x}_1^4, \tilde{x}_4^5, \tilde{x}_4^6, \tilde{x}_4^7, \tilde{x}_4^8, \tilde{x}_4^9, \tilde{x}_1^{10} . \quad (6-3)$$

It is not clear why only members 1 and 4 of the subensembles  $\tilde{x}_i^n$ ,  $i = 1, \dots, 4$  produce the flow sequence that best satisfies the vorticity equation, especially since it is these members that conform poorest to the energy spectrum slope that we expect (see column 3 of Table 2). Perhaps it is the unrealistic simplicity of the problem that we have posed that divorces it from the structure of the wind fields that are actually observed. In any case we have much to learn about this new method of producing ensembles of atmospheric flow.

Since the one flow sequence (6-3) is sufficient to demonstrate the remaining parts of the optimization procedure, we will not generate additional members of the space-time ensemble. The procedure is relatively

straightforward, using the penalty function values  $\epsilon_{ij}^n$  as the guide in assembling flow sequences whose global penalty is small.

### 6.3 Estimation of the dispersion kernels $P_S$ and $P_L$

In this problem we will take the short-term averaging interval  $T_S$  to be 1 hour and the long-term averaging time  $T_L$  to be 10 hours, i.e., the entire duration of the simulation.

In Section 2 we described the way in which a flow field description, in this case (6-3), can be used to compute the dispersion kernels  $P_S(t; \underline{x}_{S1}, \underline{x}_{S2}, \dots, \underline{x}_{SN})$  and  $P_L(\underline{x}_{S1}, \underline{x}_{S2}, \dots, \underline{x}_{SN})$  which give estimates of the short and long-term averaged concentrations through Eqs. 2-14 and 2-15. In the present problem we are concerned with sites for only two power plants, i.e.,  $\underline{x}_{S1}$  and  $\underline{x}_{S2}$ , which we will assume are confined to the "feasible" areas

$$10 < x_{S1} < 64 \quad (6-4)$$

$$5 < y_{S1} < 25$$

and

$$64 < x_{S2} < 100 \quad (6-5)$$

$$5 < y_{S2} < 25.$$

We showed in Eq. (2-16) how the kernel  $P_S$  can be represented as a cubic polynomial in the variables  $(x_{S1}, y_{S1}, x_{S2}, y_{S2})$  involving 19 coefficients  $a_1, a_2, \dots, a_{19}$ . Using the procedure described in Section 2 we have deduced kernels in polynomial form valid for any  $(\underline{x}_{S1}, \underline{x}_{S2})$  in the feasible spaces (6-4 and 6-5). (Keep in mind that both  $P_S$  and  $P_L$  give the maximum concentration value within the entire simulation region and therefore they are not functions of receptor coordinates.) For example, the vector of coefficients  $a_i, i = 1, \dots, 19$ , that describes  $P_S(t_7; \underline{x}_{S1}, \underline{x}_{S2})$  at hour  $t_7 = 7$  is

$$\tilde{a} = \begin{pmatrix} 274900.0, & -1620.0, & 1993.0, & -5678.0 \\ -491.0, & 34.52, & 68.4, & -224.5 \\ 38.55, & 143.8, & -428.8, & -0.124 \\ -0.45, & -1.746, & 5.801, & -0.057 \\ -0.853, & 0.059, & 9.62) \end{pmatrix}$$

and the corresponding vector for  $P_L(\tilde{x}_{S1}, \tilde{x}_{S2})$  is

$$\tilde{a} = \begin{pmatrix} -2200339.0 & -9252.6, & -82681.3 \\ 110164.1, & -18716.1, & 310.77 \\ -5.62, & 6349.86, & -878.45 \\ -3670.37, & 11763.7, & -3.01 \\ 1.596, & -4.388, & -137.48 & 1.21 \\ 24.13, & 4.996, & -269.45) \end{pmatrix}$$

The kernels  $P_S$  and  $P_L$  in polynomial form are used in the final step of the optimization process, namely the game theory operations, to determine the optimal locations  $(\tilde{x}_{S1}^{**}, \tilde{x}_{S2}^{**})$  of the two power plants.

#### 6.4 Determining the optimal plant sites using game theory

Referring to Section 5, Eq. (5-1), we are looking for the vector  $\tilde{z}$  of design variables

$$\tilde{z} = \begin{bmatrix} x_{S1} \\ y_{S1} \\ x_{S2} \\ y_{S2} \\ x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \end{bmatrix} \quad (6-6)$$

that results in optimal satisfaction of our cost and pollutant concentration minimization objectives while maintaining compliance with all system constraints. (In (6-6),  $x_{ij}$  is the fuel flow rate between coal mine  $i$  and power plant  $j$ .) Following is a list of the assumed values for each of the system parameters:

| Coal Mine #1                                     | Coal Mine #2                         |
|--------------------------------------------------|--------------------------------------|
| Coordinates $(\bar{x}_1, \bar{y}_1) = (40, -20)$ | $(\bar{x}_2, \bar{y}_2) = (50, -35)$ |
| Fuel cost $c^f$ (see Eq. (5-9)):                 |                                      |
| $c_1^f = \$0.40/\text{kg}$                       | $c_2^f = \$0.60/\text{kg}$           |
| Sulfur content of coal                           |                                      |
| $S_1 = 1\%$                                      | $S_2 = 3\%$                          |
| Ash content of coal                              |                                      |
| $a_1 = 30\%$                                     | $a_2 = 15\%$                         |
| Calorific value of coal                          |                                      |
| $k_1 = 4000 \text{ kcal/kg}$                     | $k_2 = 5000 \text{ kcal/kg}$         |
| Maximum capacity of mine                         |                                      |
| $m_1 = 95 \text{ ton/hr}$                        | $m_2 = 95 \text{ ton/hr}$            |

The assumed values of other system constraints include:

maximum fuel flow rate between mine  $i$   
and power plant  $j$   $= d_{ij} = 45 \text{ ton/hr}$  (all  $i, j$ );

thermal energy requirement of  
plant  $i$   $r_1 = r_2 = 2.88 \cdot 10^8 \text{ kcal/hr}$

fraction of fly ash leaving  
the furnace at plant  $i$  (see Eq. 5-15)  $f_1 = f_2 = 80\%$

efficiency of particle precipitator  
at plant  $i$  (see Eq. 5-15)  $E_1 = E_2 = 99.4\%$ .

The first step in finding the optimal design vector  $\underline{z}$  is to evaluate the matrix  $[M]$ , given by (5-18), whose elements are the five objective functions  $F_i(\underline{z})$ ,  $i = 1, \dots, 5$ , listed in Section 5 (see, for example, 5-8) evaluated for the  $\underline{z}_i^*$ ,  $i = 1, \dots, 5$ , which represent solutions of the extreme problems (see 5-17). For example,  $\underline{z}_1^*$  is the optimal design vector for the situation where objective function  $F_1(\underline{z})$  is minimized without regard for any of the other four objectives. Tables 3-6 list the optimal vectors  $\underline{z}_i^*$  and the associated value of the objective function,  $F_i(\underline{z}_i^*)$ , which were obtained for different initial guesses of  $\underline{z}_i^*$ . The point to note is that in optimization problems one is never assured of having found the global minimum of the objective function. One always finds local minima near the starting value

of the design vector, so it is necessary to consider a number of widely scattered initial points to arrive at an estimate of the global optimum design. (For brevity we have omitted details on  $\tilde{z}_5^*$ , which pertains to the extreme problem in which the only objective is minimizing long-term particulate concentration.)

Table 3. Optimum design vector  $\tilde{z}_1^*$  for the extreme problem in which the only objective is the minimization of cost. The optimal vector  $\tilde{z}_1^*$  and the associated value of the objective function,  $F_1(\tilde{z}_1^*)$ , are given for two different starting points  $\tilde{z}_1$  of the search procedure.

| Initial design vector                                                                       | Optimum design vector                                                                          |
|---------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------|
| $\tilde{z}_1 = (32.5, 17.5, 90.0, 15.0, 30.0, 30.0, 40.0, 40.0); F_1(\tilde{z}_1) = 152509$ | $\tilde{z}_1^* = (42.6, 10.0, 80.0, 5.0, 44.4, 44.3, 22.1, 22.1); F_1(\tilde{z}_1^*) = 119731$ |
| $\tilde{z}_1 = (12.0, 22.0, 98.0, 22.0, 20.0, 40.0, 44.0, 30.0); F_1(\tilde{z}_1) = 162686$ | $\tilde{z}_1^* = (41.6, 10.2, 80.5, 5.0, 24.2, 33.2, 38.2, 31.1); F_1(\tilde{z}_1^*) = 124823$ |

Table 4. Optimum design vector  $\tilde{z}_2^*$  for the extreme problem in which the only objective is minimizing short term  $S_{O2}$  concentration. The optimal vector  $\tilde{z}_2^*$  and the associated value of the objective function  $F_2(\tilde{z}_2^*)$  are given for three different starting points  $\tilde{z}_2$  of the search procedure.

| Initial design vector                                                                      | Optimum design vector                                                                           |
|--------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|
| $\tilde{z}_2 = (32.5, 17.5, 90.0, 15.0, 30.0, 30.0, 40.0, 40.0); F_2(\tilde{z}_2) = 76636$ | $\tilde{z}_2^* = (22.0, 16.5, 99.3, 14.8, 44.96, 45.0, 21.6, 21.6); F_2(\tilde{z}_2^*) = 51674$ |
| $\tilde{z}_2 = (50.0, 20.0, 85.0, 22.0, 40.0, 44.0, 30.0, 24.0); F_2(\tilde{z}_2) = 75354$ | $\tilde{z}_2^* = (19.6, 16.8, 98.0, 14.5, 45.0, 45.0, 21.6, 21.6); F_2(\tilde{z}_2^*) = 51771$  |
| $\tilde{z}_2 = (12.0, 12.0, 98.0, 24.0, 20.0, 25.0, 44.0, 40.0); F_2(\tilde{z}_2) = 78611$ | $\tilde{z}_2^* = (15.8, 17.2, 99.8, 15.0, 20.7, 25.6, 41.1, 38.0); F_2(\tilde{z}_2^*) = 67343$  |

Table 5. Optimum design vector  $\bar{z}_3^*$  for the extreme problem in which the only objective is minimizing long term  $\text{SO}_2$  concentration. The optimal vector  $\bar{z}_3^*$  and the associated value of the objective function  $F_3(\bar{z}_3)$  are given for three different initial points  $\bar{z}_3$  of the search procedure.

| Initial design vector                                                                    | Optimum design vector                                                                        |
|------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------|
| $\bar{z}_3 = (32.5, 17.5, 90.0, 15.0, 30.0, 30.0, 40.0, 40.0); F_3(\bar{z}_3) = 1120455$ | $\bar{z}_3^* = (20.3, 10.1, 100.0, 7.8, 44.8, 44.9, 23.1, 22.9); F_3(\bar{z}_3^*) = 383186$  |
| $\bar{z}_3 = (50.0, 20.0, 85.0, 22.0, 40.0, 44.0, 30.0, 24.0); F_3(\bar{z}_3) = 1162334$ | $\bar{z}_3^* = (49.4, 19.5, 83.6, 25.0, 41.6, 45.0, 24.7, 21.96); F_3(\bar{z}_3^*) = 950639$ |
| $\bar{z}_3 = (12.0, 12.0, 98.0, 24.0, 20.0, 25.0, 44.0, 40.0); F_3(\bar{z}_3) = 1156838$ | $\bar{z}_3^* = (21.1, 10.2, 80.9, 25.0, 44.9, 44.5, 21.8, 22.1); F_3(\bar{z}_3^*) = 769642$  |

Table 6. Optimum design vector  $\bar{z}_4^*$  for the extreme problem in which the only objective is minimizing short term particulate concentration. The optimal vector  $\bar{z}_4^*$  and the associated value of the objective function  $F_4(\bar{z}_4)$  are given for three different initial points  $\bar{z}_4$  of the search procedure.

| Initial design vector                                                                 | Optimum design vector                                                                      |
|---------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------|
| $\bar{z}_4 = (32.5, 17.5, 90.0, 15.0, 30.0, 30.0, 40.0, 40.0); F_4(\bar{z}_4) = 1839$ | $\bar{z}_4^* = (20.3, 15.4, 100.0, 15.9, 17.6, 17.6, 43.5, 43.7); F_4(\bar{z}_4^*) = 1327$ |
| $\bar{z}_4 = (32.5, 17.5, 90.0, 15.0, 40.0, 40.0, 30.0, 30.0); F_4(\bar{z}_4) = 2023$ | $\bar{z}_4^* = (21.1, 18.2, 99.1, 13.5, 16.8, 16.1, 44.1, 44.8); F_4(\bar{z}_4^*) = 1317$  |
| $\bar{z}_4 = (50.0, 20.0, 85.0, 22.0, 40.0, 44.0, 30.0, 24.0); F_4(\bar{z}_4) = 2448$ | $\bar{z}_4^* = (43.6, 25.0, 86.5, 18.5, 34.2, 40.1, 30.2, 25.5); F_4(\bar{z}_4^*) = 1983$  |

Using the optimal vectors  $\underline{z}_i^*$  found for each of the extreme problems, we form the matrix  $\underline{M}$  given by (5-18). We get

$$[\underline{M}] = \begin{bmatrix} 119690 & 70931 & 862059 & 2557 & 31076 \\ 141600 & 51619 & 677399 & 1888 & 24782 \\ 139291 & 61683 & 382997 & 2198 & 13649 \\ 143912 & 70884 & 874039 & 1318 & 16247 \\ 152780 & 82455 & 522618 & 1855 & 11758 \end{bmatrix}$$

Next we form the super criterion  $S(\underline{w})$  defined by (5-19):

$$S(\underline{w}) = \frac{152780}{F_1(\underline{z}_c^*)} \cdot \frac{82455}{F_2(\underline{z}_c^*)} \cdot \frac{874039}{F_3(\underline{z}_c^*)} \cdot \frac{2557}{F_4(\underline{z}_c^*)} \cdot \frac{31076}{F_5(\underline{z}_c^*)}$$

in which  $\underline{z}_c^*$  is the vector that minimizes the combined objective

$$F_c(\underline{z}) = w_1 \frac{F_1(\underline{z})}{119690} + w_2 \frac{F_2(\underline{z})}{51619} + w_3 \frac{F_3(\underline{z})}{382997} + w_4 \frac{F_4(\underline{z})}{1318} + w_5 \frac{F_5(\underline{z})}{11758}$$

subject to constraints (5-2, 3, 4, 5). Following the optimal search procedure outlined in Section 5, we arrive finally at the pareto-optimal design  $\underline{z}_c^*$  given in Table 7.

Table 7. Pareto-optimal design vector  $\tilde{z}_C^*$  obtained using the game theory concepts presented in Section 5.

| Initial values                                                          | Final, optimal, values                                                                                                                                                                       |
|-------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| $\tilde{w} = 0.1312 : 0.1626 : 0.2357 : 0.1997 : 0.2707$                | $\tilde{w}^* = 0.146 : 0.1945 : 0.193 : 0.265933 : 0.2095$                                                                                                                                   |
| $\tilde{z}_C = (32.5, 17.5, 90.0, 15.0, 30.0, 30.0, 40.0, 40.0)$        | $\tilde{z}_C^* = (32.5, 17.5, 90.0, 15.0, 30.0, 30.0, 40.0, 40.0)$                                                                                                                           |
| $\tilde{z}_C^* = (29.26, 12.62, 98.63, 8.05, 38.43, 38.38, 39.8, 39.8)$ | $\tilde{z}_C^* = (24.47, 11.68, 100.0, 7.53, 34.88, 34.73, 35.04, 35.08)$                                                                                                                    |
| $S(\tilde{w}) = 2.66$                                                   | $S(\tilde{w}) = 5.988$ with<br>$F_1(\tilde{z}_C^*) = 150713$<br>$F_2(\tilde{z}_C^*) = 75020$<br>$F_3(\tilde{z}_C^*) = 487562$<br>$F_4(\tilde{z}_C^*) = 2019$<br>$F_5(\tilde{z}_C^*) = 13124$ |

Recall that the elements of the matrix  $[M]$  given above are the objective function values  $F_i(\tilde{z}_j^*)$ , which represent the degree to which objective  $i$  is satisfied in the situation where an attempt was made to meet only objective  $j$ . Comparing those "noncooperative game" results with the objective function values given in Table 7 for the outcome of the cooperative optimization effort, we see that the effort to satisfy the five objectives jointly has resulted in design values that satisfy each objective to a degree intermediate between the worst and best that would result if the objectives were treated individually. It appears that using the starting vector  $\tilde{w}$  given in Table 7, the pareto-optimal outcome favors objectives 4 and 5 more than 1 and 2 in that the objective function

$F_4(\underline{z}_C^*)$  and  $F_5(\underline{z}_C^*)$  given in the table are closer to their absolute minimum values, given by the corresponding diagonal elements of  $[M]$  above, than the pareto-optimal values of  $F_1$  and  $F_2$ . Finding the global optimum vector  $\underline{z}^*$  requires that the process summarized in Table 7 be repeated for different initial values of the weighting vector  $\underline{w}$ .

## 7. Conclusion

This study applies methods of operations research to two aspects of air pollution modeling: (1) the generation of wind fields for use in models of regional scale transport, diffusion and chemical reaction; and (2) the application of diffusion models in studies of optimal pollution control strategies.

The newly developed method of producing wind fields differs fundamentally from existing methods in that it yields ensembles of flows rather than a single flow specification. In this regard it is consistent with the fact that physical laws and discrete meteorological data do not uniquely specify the wind field.

We demonstrated, using a 2-D fluid as an example, that a set of  $J$  discrete observations of the fluid velocity at arbitrary sites at a given instant and the equation of mass continuity together define an infinite set of functions any one of which is a possible description of the flow that existed at the time the measurements were made. We hypothesized that of these functions the ones most likely to describe the actual flow are those whose kinetic energy spectra conform to given empirical specifications. An optimization procedure was described that allows one to select functions possessing the desired spectral characteristics

from the infinite set of possible flows. Using the sets of probable flows derived for each hour that wind data are available, we constructed sequences of flow fields to represent the evolution of velocity in time. We proposed that the probability of any given sequence of velocity fields is proportional to the degree to which that sequence satisfies the principle of momentum conservation. We showed that dynamic programming can be used to select from the vast set of possible flow sequences a finite subset containing the most probable of all sequences. This subset constitutes an ensemble of flow fields in space-time whose members map one-to-one onto a corresponding ensemble of species concentration distributions. Our underlying philosophy is that it is only properties of the concentration ensemble, rather than properties of any one of its members, that constitute a proper basis for decision making in applied air pollution modeling. Because it is only in the ensemble properties that the inherent uncertainty in specifying atmospheric flow is taken into account.

The newly developed method of model application allows one to find solutions to air quality problems that result in optimal compliance with multiple objectives subject to given constraints. The objectives we considered are minimization of cost, including the cost of operating pollution control equipment; and the minimization of the peak and long-term averaged concentrations of each of two air pollutants -- a total of five objectives in all. One of the useful functions that a technique of this type can provide is the ability to assess the penalties associated with efforts to satisfy one or several objectives at the exclusion of others. For example, how much increase in long-term averaged concen-

tration would result if only operating cost and peak concentration are minimized jointly? Or, what sacrifice in air quality would result if pollutant concentrations were treated as an objective to be minimized jointly with cost rather than a rigid system constraint? These are questions of potentially far reaching significance which we hope to pursue in a future study.

Although the technique that we have developed has its merits, it also has drawbacks. One of these is a problem common to all optimization procedures, namely, the inability to determine that a given solution is the global optimum. The method produces solutions that are optimum within some local region of the parameter space; but to find the global minimum solution, one must generally apply the method to a number of different starting points in the parameter space and select from the resulting set of local optimum solutions the one that best satisfies the objective function. This procedure can be expensive computationally, especially if the air quality objectives are expressed in terms of a model or models consisting of differential equations. For this reason, we believe that the only feasible applications of our technique, with present generation computers, are in problems where the air quality objectives are expressible either in closed functional forms or as matrices of discrete values.

## References

- Nastrom, G. D., K. S. Gage and W. H. Jasperson (1984): Kinetic energy spectrum of large- and mesoscale atmospheric processes. Nature, Vol. 310, pp. 36-38.
- Lamb, R. G. (1984): Air pollution models as descriptors of cause-effect relationships. Atmos. Environ., Vol. 18:3, pp. 591-606.
- (1983): A regional scale (1000 km) model of photochemical air pollution. Part 1. Theoretical formulation. EPA report EPA-600/3-83-035. 238 pages.
- (1976) (L76): The Calculation of Long-Term Atmospheric Pollution Concentration Statistics Using the Concept of a Macro-Turbulence. Preprint volume, Third Symposium on Atmospheric Turbulence, Diffusion and Air Quality. American Meteorological Society. 553-560.
- Seinfeld, J. H. and C. P. Kyan (1971): Determination of Optimal Air Pollution Control Strategies. Socio-Econ. Plan. Sci. Vol. 5, 173-190.
- Kyan, C. P. and J. H. Seinfeld (1972): Determination of Optimal Multiyear Air Pollution Control Policies. J. Dynamic. Sys. Meas. and Control Vol. 94G, 266-274.